

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:39 ; Search time 1836.07 Seconds

(Without alignments)  
12110.869 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373

Sequence: 1 atgcattgttgaattga.....gacaaatgttgaataatg 1373

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	765.8	55.8	834	13	BI766766 603056866
2	753.2	54.9	948	14	BO707185 AGENCOURT
3	747.2	54.4	963	14	BO671259 AGENCOURT
4	743.4	54.1	777	13	BI819200 603034614
5	725.4	52.8	1071	14	BM921213 AGENCOURT
6	701.4	51.1	731	13	BI871711 603395825

7	688	50.1	828	13	BI596681 603243254
8	677.4	49.3	728	13	BI870393 603395641
9	584.4	42.6	1027	13	BI966060 1672804.X
10	581.4	42.3	1027	17	AF163779 AF163779
11	562	40.9	940	14	BO884231 AGENCOURT
12	534.6	38.9	568	14	BM971606 UI-CE-ECL
13	531.6	38.7	910	12	BI110063 602279667
14	516.8	37.6	531	13	BI824443 603038693
15	493.6	36.0	918	12	BF577781 602092080
16	458.2	33.4	1033	11	AK020909 Mus muscu
17	454.4	33.1	561	10	AW763237 UR70409.Y
18	444	32.3	951	14	BO674188 AGENCOURT
19	442.8	32.3	587	12	BG668319 602638232
20	440.2	32.1	456	13	BI966255 1672804.Y
21	436.6	31.8	440	13	BM128059 1672804.Y
22	431.6	31.4	538	12	BF821434 MRI-RT003
23	404	29.4	436	17	AO890280 HS_3188.B
24	399.4	29.1	785	13	BI762908 603047966
25	394	28.7	413	9	AI422796 tcf22e06.x
26	389.2	28.3	698	13	BI906850 603064633
27	374.2	27.3	416	9	AI291866 qm6c02.x
28	372.6	27.1	584	10	AW917574 EST348878
29	365.4	26.6	422	13	BI677255 1687402.X
30	365	26.6	399	9	AI913541 wa12c01.x
31	363.4	26.5	367	17	AO100365 HS_3054.A
32	359	26.1	407	9	BI965174 1634607.Y
33	358.4	26.1	474	9	AI221985 q995a10.x
34	357.6	26.0	894	13	BI908274 603068526
35	352.8	25.7	397	10	AW131279 xF61404.X
36	347.2	25.3	409	9	AI669243 wc13901.X
37	342.6	25.0	710	12	BE858778 7E95005.X
38	341	24.8	377	13	BM509016 1115005.Y
39	339.8	24.7	372	13	BI677256 1687402.Y
40	330	24.0	374	13	BM505649 1115005.X
41	328	23.9	367	13	BM662622 UI-E-CO1
42	327.8	23.9	650	12	BG404836 602420160
43	327	23.8	346	12	BE858822 7G38604.X
44	327	23.8	471	9	AA221610 my18d09.X
45	326	23.7	337	10	AW150534 xm45912.X

## ALIGNMENTS

RESULT 1  
BI766766  
LOCUS  
DEFINITION  
603056866F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5406217 5',  
mRNA sequence.  
ACCESSION  
BI766766  
VERSION  
BI766766.1 GI:15758344  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 834)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: gcaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L14M1517 row: C column: 18  
High quality sequence stop: 772.  
Location/Qualifiers  
1..834

FEATURES  
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5206217"
/clone_lib="NIH_MGC_112"
/lab_host="DH10B"
/Note="Organ: Pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

BASE COUNT      154 a      287 c      223 g      170 t
ORIGIN
Query Match      55.8%; Score 765.8; DB 13; Length 834;
Best Local Similarity 98.0%; Pred. No. 8.1e-157;
Matches 818; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

QY 272 CCCAGAGAGAGCTGTGGCGAGAGAGACCCGTCGGAACTGAATCCCAAGAC 331
DB 1 CCCAGAGAGAGCTGTGGCGAGAGAGACCCGTCGGAACTGAATCCCAAGAC 60
QY 332 AAGAAAGCCAGATCTCTGCGCTTCTCCTGAACGAGTACTTCCGCTCCGAGAAGTGAC 391
DB 61 AAGAAAGCCAGATCTCTGCGCTTCTCCTGAACGAGTACTTCCGCTCCGAGAAGTGAC 120
QY 392 CTAAAGCGCGAAACACGGGCTCGAAGAGCATCCGAGCCATATGAAGTTATCCAC 451
DB 121 CTAAAGCGCGAAACACGGGCTCGAAGAGCATCCGAGCCATATGAAGTTATCCAC 180
QY 452 GACC-TGGACAGAGAGAGAGAGAGAGAGTGTGGAGGAGCACTGATGCTGGAGGAA 510
DB 181 GACCAGAGAGAGAGAGAGAGAGAGTGTGGAGGAGCACTGATGCTGGAGGAA 240
QY 511 GCCAGATCAACAGCTCAACGCTCTGCGCTACACCGCCAGATCGGGAGTTATAGTC 570
DB 241 GCCAGATCAACAGCTCAACGCTCTGCGCTACACCGCCAGATCGGGAGTTATAGTC 300
QY 571 ACCCGGCTGGGCTCTACTACTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
DB 301 ACCCGGCTGGGCTCTACTACTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 631 TACCTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
DB 361 TACCTGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 691 TCAGCCACTGCGGCGCAGTTCCTCGGGGCCACGCTCCGCTCTGCCAGATGCTGGGCTG 750
DB 421 TCAGCCACTGCGGCGCAGTTCCTCGGGGCCACGCTCCGCTCTGCCAGATGCTGGGCTG 480
QY 751 TTGGCCCTGGGCGAGGCTCTCTGCGGATCCGCACTCCCTCGGGGCCATCTCAAG 810
DB 481 TTGGCCCTGGGCGCA-GGTCCTCTCGCGGATCCGCACTCCCTCGGGGCCATCTCAAG 539
QY 811 GTGGCCCTCTCTCACTACTCTCGGACTCTTCAGAGTTACTGAGGGGCCCTGGTCTCC 870
DB 540 GTGGCCCTCTCTCACTACTCTCGGACTCTTCAGAGTTACTGAGGGGCCCTGGTCTCC 599
QY 871 CCACAGTGTGCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCCGCTGCC-TC 929
DB 600 CCACAGTGTGCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCCGCTGCCCTTC 659
QY 930 TGGCCCAACCTCAAGCGCTCTTGTCTCCAGACCTGCTCCCTCTAGAGAGCTGCTGG 989
DB 660 TGGCCCAACCTCAAGCGCTCTTGTCTCCAGACCTGCTCCCTCTAGAGAGCTGCTGG 719
QY 990 CCGTGTCAAGTGTTCCTCAATCCCAATATAGATATTCCTTATCTTACAAT- 1048
DB 720 CCGTGTCAAGTGTTCCTCAATCCCAATATAGATATTCCTTATCTTAAAAAT 779

```

```

QY 1049 -CCCCACGCGCCACTCTCCACTCAGTACCTCCCAATCTGACCTTTGAG 1102
DB 780 CCCCCAAGCGCCACTCTCCACTCAGTACCTCCCAATCTGACCTTTGAG 834

RESULT 2
LOCUS      BQ707185
DEFINITION AGENCOURT_8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6278608
5', mRNA sequence.
ACCESSION BQ707185
VERSION   BQ707185.1 GI:21846084
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapds-remail.nih.gov
            Tissue Procurement: Dr. Mark Watson
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM2466 row: n column: 17
            High quality sequence start: 24
            High quality sequence stop: 550.
            Location/Qualifiers
                1..948
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6278608"
                    /clone_lib="NIH_MGC_113"
                    /lab_host="DH10B (phage-resistant)"
                    /Note="Organ: spleen; Vector: pOT7; Site_1: XhoI; Site_2:
                    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                    into EcoRI/XhoI sites using the following 5' adaptor:
                    GGCACAG(G). Library constructed by Ling Hong in the
                    laboratory of Gerald M. Rubin (University of California,
                    Berkeley) using zap-cDNA synthesis kit (Stratagene) and
                    Superscript II RT (Life Technologies). Note: this is a
                    NIH_MGC Library."

BASE COUNT      173 a      334 c      252 g      187 t      2 others
ORIGIN
Query Match      54.9%; Score 753.2; DB 14; Length 948;
Best Local Similarity 95.6%; Pred. No. 4.7e-154;
Matches 807; Conservative 0; Mismatches 30; Indels 7; Gaps 3;

QY 305 ACCCGTCGAACTGAATCCCAAGAGAAAGAGATCCTGCGCTTCTGTAAC 364
DB 1 ACCCGTCGAACTGAATCCCAAGAGAAAGAGATCCTGCGCTTCTGTAAC 60
QY 365 GACTAGTTGGGCTCCGAGAGTGCACCTTAAGGCGGAAACACGGGCTCGAAGCGCA 424
DB 61 GACTAGTTGGGCTCCGAGAGTGCACCTTAAGGCGGAAACACGGGCTCGAAGCGCA 120
QY 425 TGGCAGCCATTTATGAAGTTCAACAGACTGTGAGAGAGAGAGAGAGAGAGAGAG 484
DB 121 TGGCAGCCATTTATGAAGTTCAACAGACTGTGAGAGAGAGAGAGAGAGAGAGAG 176
QY 485 ACGGACAGTGAATGCTGGGAGAGAGAGAGATCAAGAGCTCCAGCGCTGCTACA 544
DB 177 ACGGACAGTGAATGCTGGGAGAGAGAGAGATCAAGAGCTCCAGCGCTGCTACA 236
QY 545 ACCGACAGATCGGGAGTTATAGTACCCCGGCTGGCTTACTACTGTACAG 604

```





FEATURES	source
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	Email: cgsb@remail.nih.gov
ORGANISM	Human.
LOCUS	B1819200
DEFINITION	777 bp mRNA linear EST 04-OCT-2001
ACCESSION	B1819200
VERSION	B1819200.1
KEYWORDS	EST.
SOURCE	Human.
ORGANISM	Human.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 777)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
ORGANISM	Contact: Robert Strausberg, Ph.D.
LOCUS	Email: cgsb@remail.nih.gov
DEFINITION	Tissue Procurement: Life Technologies, Inc.
ACCESSION	cDNA Library Preparation: Life Technologies, Inc.
VERSION	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
KEYWORDS	DNA Sequencing by: Incyte Genomics, Inc.
SOURCE	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
ORGANISM	Plate: L18M1437 row: 1 column: 03
REFERENCE	High quality sequence stop: 759.
AUTHORS	Location/Qualifiers
TITLE	1. 777
JOURNAL	/organism="Homo sapiens"
COMMENT	/db_xref="taxon:9606"
ORGANISM	/clone="IMAGE:5175698"
LOCUS	/clone_lib="NIH-MGC_115"
DEFINITION	/lab_host="DH10B"
ACCESSION	/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."
VERSION	021. Note: this is a NIH-MGC Library."
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORG	

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752561"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      226 a      346 c      279 g      220 t
ORIGIN

Query Match      52.8%; Score 725.4; DB 14; Length 1071;
Best Local Similarity 93.4%; Pred. No. 5.6e-148;
Matches 836; Conservative 0; Mismatches 46; Indels 13; Gaps 7;

QY 412 GCTCGAAGAGCATGCGACCCATTATGATTCACGACCTGGACAGAGGAGCG 471
    |||
DB 12 GCTCGAAGAGCATGCGACCCATTATGATTCACGACCTGGACAGAGGAGCG 71
    |||

QY 472 CAGGAGGTGTGGAGGAGCATGAGTGGTGGAGGAGGAGGAGGAGGAGGAGG 531
    |||
DB 72 CAGGAGGTGTGGAGGAGCATGAGTGGTGGAGGAGGAGGAGGAGGAGGAGG 131
    |||

QY 533 CCTCTGCGCTACACCGCCAGATCGGGAGTTTATGATCACCGGGGCTGGCTACTAC 591
    |||
DB 133 CCTCTGCGCTACACCGCCAGATCGGGAGTTTATGATCACCGGGGCTGGCTACTAC 191
    |||

QY 592 CTGTACTGTACAGTGCATTTATGATGAGGAGGAGGAGGAGGAGGAGGAGG 651
    |||
DB 192 CTGTACTGTACAGTGCATTTATGATGAGGAGGAGGAGGAGGAGGAGGAGG 251
    |||

QY 653 GGGAGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 711
    |||
DB 252 GGGAGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 311
    |||

QY 712 CTGGGGCCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGGCTGGGCGCAGGCTCC 771
    |||
DB 312 CTGGGGCCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGGCTGGGCGCAGGCTCC 371
    |||

QY 772 TCCCTGCGGATCCGACCCCTCCCTGGGCGCATCTCAGAGGCTGCCCTTCTCAGCTAC 831
    |||
DB 372 TCCCTGCGGATCCGACCCCTCCCTGGGCGCATCTCAGAGGCTGCCCTTCTCAGCTAC 431
    |||

QY 832 TTGGGACTCTTCAGGTTTACTGAGGAGGCGCGTGTCTCCCAAGTCGTCAGAGGCTGGC 891
    |||
DB 432 TTGGGACTCTTCAGGTTTACTGAGGAGGCGCGTGTCTCCCAAGTCGTCAGAGGCTGGC 491
    |||

QY 892 GGCCTCCCTCGACAGCTCTCTGGGCAACCGGTCCTCCCTGCGCCACCTTCAGCCGCTCTT 951
    |||
DB 492 GGCCTCCCTCGACAGCTCTCTGGGCAACCGGTCCTCCCTGCGCCACCTTCAGCCGCTCTT 551
    |||

QY 952 TCCCTCAGAGCTGGCCCTTCCCTCTAGAGGCTGCTGGGCTGTCTACAGTGTTCATCC 1011
    |||
DB 552 TCCCTCAGAGCTGGCCCTTCCCTCTAGAGGCTGCTGGGCTGTCTACAGTGTTCATCC 611
    |||

QY 1012 CACATTAATATACAGTATTCCTCTTATCTTAACTCCCGACCGCCACTCTCAGCT 1071
    |||
DB 612 CACATTAATATACAGTATTCCTCTTATCTTAACTCCCGACCGCCACTCTCAGCT 671
    |||

QY 1072 CACTAGACTCCCAATCCCTGACCTTTGAGGCCCGCCAGTATCTGACT-CCGCCCTGGC 1130
    |||
DB 672 CACTAGACTCCCAATCCCTGACCTTTGAGGCCCGCCAGTATCTGACT-CCGCCCTGGC 731
    |||

QY 1131 CACAGACCCCC-AGGGCATGTG-TTCACTGTACTCTGTGGGCAAGGA--TGGTCTCAAA 1186
    |||
DB 732 CACAGACCCCCCAAGGCAATGTGTTCACCTGTGTGGGGCAAGGAATGGGTCCAAAGA 791
    |||

```

```

QY 1187 AGACCCCACTTCAGG---CACTAAGAGGGGCTGACCTTGGGGCAGG---AAGCCAAG 1239
    |||
DB 792 AAACCCCACTTCGCGGACCTAAAGGGGCTGGGACCTGGGGGCAAGAAAGCCCAAG 851
    |||

QY 1240 AGACTGGG-CTTAGGCGCAGAGTGTCCCAATGTGAGGGGCGAGAAACAGACAG 1293
    |||
DB 852 AAACCTGGGCTTAGGCGCAGAGGAGTGTCCCAATGTGAGGGGCGAGAAACAGAG 906
    |||

RESULT 6
BI871711
LOCUS 603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 731)
AUTHORS
NIH-MGC http://mgc.nhl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12034 row: b column: 07
High quality sequence stop: 728.
Location/Qualifiers
source
1. 731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_lib="NIH_MGC_90"
/tissue="adenoecarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: liver; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      125 a      242 c      228 g      136 t
ORIGIN

Query Match      51.1%; Score 701.4; DB 13; Length 731;
Best Local Similarity 99.3%; Pred. No. 9e-143;
Matches 726; Conservative 0; Mismatches 1; Indels 4; Gaps 2;

QY 192 CCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251
    |||
DB 1 CCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
    |||

QY 252 GCTGTCCGCC---CAGAGACCTGCGCAGAGAGCTGCTGGGAGAGAGAGCAGAGACC 308
    |||
DB 61 GCTGTCCGCCAGCAGAGACCTGCGCAGAGAGCTGCTGGGAGAGAGAGCAGAGACC 120
    |||

QY 309 GTTCGAATCTGAATCCCGACAGAAAGAGCAGATCTGCGCTTCTGAACGACT 368
    |||
DB 121 GTTCGAATCTGAATCCCGACAGAAAGAGCAGATCTGCGCTTCTGAACGACT 180
    |||

QY 369 AGTTGGCTCTCCAGAGAGTGCATTAAGCCGGGAGAAACAGGGCTCCGAAGAGCATGCC 428
    |||
DB 181 AGTTGGCTCTCCAGAGAGTGCATTAAGCCGGGAGAAACAGGGCTCCGAAGAGCATGCC 240
    |||

```

QY 429 AGCCATTATGAAAGTTCATCCAGACCTGACAGAGGAGGAGGAGGAGTGTGACAG 488  
 Db 241 AGCCATTATGAAAGTTCATCCAGACCTGACAGAGGAGGAGGAGGAGTGTGACAG 300  
 QY 489 GACAGTATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 548  
 Db 301 GACAGTATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
 QY 549 CCAGATGAGGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 608  
 Db 361 CCAGATGAGGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 420  
 QY 609 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 668  
 Db 421 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 QY 669 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 728  
 Db 481 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 QY 729 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788  
 Db 541 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599  
 QY 789 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848  
 Db 600 CTTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659  
 QY 849 TCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908  
 Db 660 TCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719  
 QY 909 CTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 919  
 Db 720 CTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 730

RESULT 7 828 bp mRNA linear EST 07-SEP-2001  
 B1596681  
 LOCUS 603243254F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5285892 5',  
 DEFINITION  
 mRNA sequence.  
 B1596681  
 ACCESSION  
 B1596681.1 GI:15489620  
 VERSION  
 B1596681.1  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (bases 1 to 828)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE  
 Unpublished (1999)  
 JOURNAL  
 Contact: Robert Strausberg, Ph.D.  
 COMMENT  
 Email: e9apds-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1AM1722 row: k column: 13  
 High quality sequence stop: 776.  
 Location/Qualifiers  
 1. 828  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5285892"  
 /clone\_1ib="NIH\_MGC\_96"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtc9ag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',  
 size-selected for average insert size 2.3 kb and  
 normalized to 10<sup>7</sup>. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH-MGC Library."  
 BASE COUNT 155 a 278 c 223 g 172 t  
 ORIGIN

Query Match 50.1%; Score 688; DB 13; Length 828;  
 Best Local Similarity 95.3%; Pred. No. 7.8e-140;  
 Matches 732; Conservative 0; Mismatches 30; Indels 6; Gaps 2;

QY 312 GGAAGTAAATCCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 371  
 Db 43 GGAAGTAAATCCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 102  
 QY 372 TCGGCTCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 431  
 Db 103 TCGGCTCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162  
 QY 432 CCATTATGAAAGTTCATCCAGACCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 491  
 Db 163 CCATTATGAAAGTTCATCCAGACCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 222  
 QY 492 AGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 551  
 Db 223 AGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 282  
 QY 552 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 611  
 Db 283 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402  
 QY 612 TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 671  
 Db 343 TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402  
 QY 672 GCGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 731  
 Db 403 GCGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 462  
 QY 732 CTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 791  
 Db 463 CTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 522  
 QY 792 CCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 851  
 Db 523 CCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 582  
 QY 852 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 911  
 Db 583 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 642  
 QY 912 TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 971  
 Db 643 TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 702  
 QY 972 CTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1025  
 Db 703 CTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 762  
 QY 1026 ATTCCACTCTTATCTTAACTCCAGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1073  
 Db 763 TTTCACACTCTTATCTTAACTCCAGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 810

RESULT 8 728 bp mRNA linear EST 11-OCT-2001  
 B1870393  
 LOCUS 603395641F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5405459 5',  
 DEFINITION



fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806. 0.5 microgram single stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others  
ORIGIN

Query Match 42.6%; Score 584.4; DB 13; Length 609;  
Best Local Similarity 99.7%; Pred. No. 3.1e-117;  
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 ACCCTCCCTGGGCGCCATCTCAAGGCTGCCCTCTCTCACTACTTGGACTCTTCAG 846  
DB 609 ACCCTCCCTGGGCGCCATCTCAAGGCTGCCCTCTCTCACTACTTGGACTCTTCAG 550  
QY 847 GTTCACTGAGGGGCGCCGCTCTCCCAAGCTGCTCCAGGCTCCGGCTCCCTCGACAG 906  
DB 549 GTTCACTGAGGGGCGCCGCTCTCCCAAGCTGCTCCAGGCTCCGGCTCCCTCGACAG 490  
QY 907 CTCTCTGGGACCCCGTCCCTCTGCCCCACCCCTCAAGCCGCTCTTCTCGAGACTGCC 966  
DB 489 CTCTCTGGGACCCCGTCCCTCTGCCCCACCCCTCAAGCCGCTCTTCTCGAGACTGCC 430  
QY 967 CTTCCCTCTAGAGGCTCCCTGCGGCTGTCTACAGTCTTTCATCCCACTAAATACAGTA 1026  
DB 429 CTTCCCTCTAGAGGCTCCCTGCGGCTGTCTACAGTCTTTCATCCCACTAAATACAGTA 370  
QY 1027 TTCCCACTCTTATCTTACAACTCCCAAGCCCACTCTCACTCACTAGCTCCCAAT 1086  
DB 369 TTCCCACTCTTATCTTACAACTCCCAAGCCCACTCTCACTCACTAGCTCCCAAT 310  
QY 1087 CCTGACCTCTTGAAGGCGCCCAAGTGAATCTGCCCTCGGCGACAGACCCCGAGGGC 1146  
DB 309 CCTGACCTCTTGAAGGCGCCCAAGTGAATCTGCCCTCGGCGACAGACCCCGAGGGC 250  
QY 1147 ATTGTGTTCACTGACTCTGTGGGCGAAGAGTGGTCCAGAAAGCCCACTCAAGGCACTA 1206  
DB 249 ATTGTGTTCACTGACTCTGTGGGCGAAGAGTGGTCCAGAAAGCCCACTCAAGGCACTA 190  
QY 1207 AGAGGGGCTGGACCTGGCGGCGAGAAAGCAAGACTGGGCTTAGGCCAGAGTTCCCA 1286  
DB 189 AGAGGGGCTGGACCTGGCGGCGAGAAAGCAAGACTGGGCTTAGGCCAGAGTTCCCA 130  
QY 1267 AATGTAGGGGCGGAGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAA 1326  
DB 129 AATGTAGGGGCGGAGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAA 70  
QY 1327 AACGATATATTTTATTTATTTATTTGACAAATGTGTAATATG 1373  
DB 69 AACGATATATTTTATTTATTTATTTGACAAATGTGTAATATG 23

RESULT 10  
AF163779 1027 bp DNA linear GSS 29-AUG-2000  
LOCUS AF163779 Human Homo sapiens genomic clone BAC750E14, DNA sequence.  
DEFINITION AF163779  
ACCESSION AF163779.1 GI:5726439  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
JOURNAL Cousin, P., Billotte, J., Chaubert, P., and Shaw, P. H.  
Physical map of 17p13 and the genes adjacent to p53  
Genomics 63 (1), 60-66 (2000)

MEDLINE 20130114  
COMMENT

Contact: Shaw PH  
Experimental Oncology  
Institute of Pathology  
Rue du Bugnon 25, Lausanne, VD 1011, Switzerland  
Subclone-A2R Asc-BamHI PSL180  
Class: BAC subclone.  
Location/Qualifiers  
1. 1027  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="17p"  
/clone="BAC750E14"  
/note="Vector: PMOS Blue"

FEATURES  
source

BASE COUNT 207 a 317 c 282 g 208 t 13 others  
ORIGIN

Query Match 42.3%; Score 581.4; DB 17; Length 1027;  
Best Local Similarity 98.0%; Pred. No. 1.5e-116;  
Matches 577; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 785 GCACCTCTCCCTGGGCGCCATCTCAAGGCTGCCCTCTCTCACTACTTGGACTCTTC 844  
DB 1 GCWCCCTCCCTGGGCGCCATCTCAAGGCTGCCCTCTCTCACTACTTGGACTCTTC 60  
QY 845 AGGTTCACTGAGGGGCGCCGCTCTCCCAAGCTGCTCCAGAGGCTCCGGCTCCCTCGAC 904  
DB 61 AGGTTCACTGAGGGGCGCCGCTCTCCCAAGCTGCTCCAGAGGCTCCGGCTCCCTCGAC 120  
QY 905 AGCTCTGAGGACCCGCTCCCTCTGCCCCACCCCTCAAGCCGCTCTTGTCTCAAGACTG 964  
DB 121 AGCTCTGAGGACCCGCTCCCTCTGCCCCACCCCTCAAGCCGCTCTTGTCTCAAGACTG 180  
QY 965 CCCCTCCCTAGAGGCTGCTGGGCGCTGTTACAGTCTTTCATCCCACTAAATACAG 1024  
DB 181 CCCCTCCCTAGAGGCTGCTGGGCGCTGTTACAGTCTTTCATCCCACTAAATACAG 240  
QY 1025 TATTCACACTTATCTTACAACTCCCAAGCCCACTCTCACTCACTAGCTCCCA 1084  
DB 241 TATTCACACTTATCTTACAACTCCCAAGCCCACTCTCACTCACTAGCTCCCA 300  
QY 1085 ATCCCTGACCTTTGAGGCGCCCAAGTATCTGCACTCCCTGCGCACAGACCCCGAGG 1144  
DB 301 ATCCCTGACCTTTGAGGCGCCCAAGTATCTGCACTCCCTGCGCACAGACCCCGAGG 360  
QY 1145 GCATTGTTCACTGACTCTGTGGGCGAAGAGTGGTCCAGAAAGCCCACTCAAGGCACT 1204  
DB 361 GCATTGTTCACTGACTCTGTGGGCGAAGAGTGGTCCAGAAAGCCCACTCAAGGCACT 420  
QY 1205 TAAGAGGGGCTGGACCTGGCGGCGAGAAAGCAAGACTGGGCTTAGGCCAGAGTTCC 1264  
DB 421 TAAGAGGGGCTGGACCTGGCGGCGAGAAAGCAAGACTGGGCTTAGGCCAGAGTTCC 480  
QY 1265 CAAATGTAGGGGCGGAGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTT 1324  
DB 481 CAAATGTAGGGGCGGAGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTT 540  
QY 1325 AAAACGATATTTTATTTATTTATTTGACAAATGTGTAATATG 1373  
DB 541 AAAACGATATTTTATTTATTTATTTGACAAATGTGTAATATG 589

RESULT 11  
BO884231 940 bp mRNA linear EST 16-AUG-2002  
LOCUS BO884231  
DEFINITION AGENCOURT\_8682031 lupsk1\_sclatlc\_nerve Homo sapiens cdna clone  
IMAGE:6197488 5', mRNA sequence.  
ACCESSION BO884231  
VERSION BO884231.1 GI:22276239  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: L1AM13607 row: 3 column: 17  
 High quality sequence stop: 453.

FEATURES  
 Location/Qualifiers  
 1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6197488"  
 /clone\_lib="Lupski\_sclatic\_nerve"  
 /sex="male"  
 /tissue\_type="sclatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; CDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGTCG-3' and  
 5'-GATAGTCTGATGATGAGCGGCGGCGGCTT(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN

Query Match 40.9%; Score 562; DB 14; Length 940;  
 Best Local Similarity 96.5%; Pred. NO. 2.6e-112;  
 Matches 660; Conservative 0; Mismatches 15; Indels 9; Gaps 8;

OY 301 CAGGACCCGTCGAGACTAATCCAGAGAGAAAGGAGGATCTGGCCCTTCTCG 360  
 |||||||  
 DB 1 CAGGACCCGTCGAGACTAATCCAGAGAGAAAGGAGGATCTGGCCCTTCTCG 60  
 |||||||  
 OY 361 AACGCACTAGTTCGGCTCGCAGAAAGTCCACCTAAAGCCGGAACACAGGCTCGAAGA 420  
 |||||||  
 DB 61 AACGCACTAGTTCGGCTCGCAGAAAGTCCACCTAAAGCCGGAACACAGGCTCGAAGA 120  
 |||||||  
 OY 421 GCGATCGAGCCCATTAAGAGTTTATCCAGACCTGGACAGAGCGAGCGACAGT 480  
 |||||||  
 DB 121 GCGATCGAGCCCATTAAGAGTTTATCCAGACCTGGACAGAGCGAGCGACAGT 180  
 |||||||  
 OY 481 GTGACGGGAGATGAGTGGCTGGGAGGAGGAGGAGATCAACAGTCCAGCCCTGCGC 540  
 |||||||  
 DB 181 GTGACGGGAGATGAGTGGCTGGGAGGAGGAGGAGATCAACAGTCCAGCCCTGCGC 240  
 |||||||  
 OY 541 TACAACCCGAGATCGGGAGAGTTTATAGTCAACCCGGCTGGGCTCTACTACTGATCTGT 600  
 |||||||  
 DB 241 TACAACCCGAGATCGGGAGAGTTTATAGTCAACCCGGCTGGGCTCTACTACTGATCTGT 300  
 |||||||  
 OY 601 CAGTGCACCTTTGATGAGAGGAGAGCTGTCTACTGAAAGCTTGAGCTTCTGGTGTATGCT 660  
 |||||||  
 DB 301 CAGTGCACCTTTGATGAGAGGAGAGCTGTCTACTGAAAGCTTGAGCTTCTGGTGTATGCT 360  
 |||||||  
 OY 661 GTGCTGGCCCTGCTGCTGAGAGATTCACGCACTGGGGCCAGTTCCCTGGGGCC 720  
 |||||||  
 DB 361 GTGCTGGCCCTGCTGCTGAGAGATTCACGCACTGGGGCCAGTTCCCTGGGGCC 420  
 |||||||  
 OY 721 CAGTCCGCTCTG-CGAGTGTCTGGGCTGTGGCCCTGGGGCCA-GGGTCCCTCCCTGTC 778

DB 421 CAGTCCGCTCTGCTGCTGAGAGTGTCTGGGCTGTGGCCCTGGGGCCAGGGCCCTCCCTGC 480  
 |||||||  
 OY 779 GGATCCGACCCCTCCCTGGGCGCCATCTCAAGGCTG-CGCCCTTCTCACTAC-TTCGG 836  
 |||||||  
 DB 481 GGATCCGACCCCTCCCTGGGCGCCATCTCAAGGCTGCGCCCTTCTCACTACTTTCGG 540  
 |||||||  
 OY 837 ACTCTCCAGGTTCACTAGAGGGCCCTGCTGCCCAAGTGT-CCAGGCTGCGGGC 894  
 |||||||  
 DB 541 ACTCTCCAGGTTCACTAGAGGGCCCTGCTGCCCAAGTGT-CCAGGCTGCGGGC 600  
 |||||||  
 OY 895 TCCCTCCGACAGCTCTCT-GGGCACCCTGCTCTCTGCTCCCAACC--TCAGCGCTCTT 951  
 |||||||  
 DB 601 TCCCTCCGACAGCTCTCTGGGGAACCGGTCCCTCTGCTCCCAACCCTCAGCGCTCTT 660  
 |||||||  
 OY 952 TGCTCCAGACTGCTCCCTCTCT 975  
 |||||||  
 DB 661 TGCTCCAGACTGCTCCCTCTCT 684

RESULT 12  
 BM971606/c 568 bp mRNA linear EST 21-MAR-2002  
 LOCUS  
 DEFINITION UI-CF-EC1-ab1-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens CDNA clone  
 UI-CF-EC1-ab1-p-06-0-UI 3', mRNA sequence.  
 ACCESSION BM971606  
 VERSION  
 KEYWORDS BM971606.1 GI:19589193  
 EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 568)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171

Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com)  
 The following repetitive elements were found in this CDNA  
 sequence: 1-82, >AT-rich/low-complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 Location/Qualifiers  
 1..568

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EC1-ab1-p-06-0-UI"  
 /clone\_lib="UI-CF-EC1"  
 /tissue\_type="lung"  
 /dev\_stage="adult and fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: Lung; Vector: pVT3-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: BcoR I; Site\_2: Not I;  
 UI-CF-EC1 is a normalized CDNA library containing the  
 following tissue(s): Normal lung from adult and from fetal  
 day 64, day 87, week 19 and week 42. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand CDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded CDNA was ligated to an BcoR I







Db	596	AGGGAAGCAAGAGACCTAGAGCGCTAGAGCAACAGATGTCACAAAGTGAAGCGGCAAGAAA	655
QY	1286	AAGACAAGCTCTCTCCCTTGAGAAATTCCTGTGGATTTTAA	1327
Db	656	CAAGGAAGACCTCCCATGAGGATGCGCTGAGAAATATGACAA	697
RESULT 14			
LOCUS	B1824443	531 bp	mRNA
DEFINITION	603038693p1 NIH_MGC_115	Homo sapiens	CDNA clone IMAGE:5179510 5',
ACCESSION	B1824443		
VERSION	B1824443.1	GI:15935993	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 531)		
JOURNAL	NIH-MGC <a href="http://mhc.ncl.nih.gov/">http://mhc.ncl.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <a href="http://lmnl.lnl.gov">http://lmnl.lnl.gov</a> Plate: LMNL1447 row: j column: 23 High quality sequence stop: 529.		
FEATURES			
SOURCE	location/Qualifiers		
	1..531		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="5179510"		
	/clone_11b="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."		
BASE COUNT	72 a	223 c	119 g
ORIGIN			117 t
Query Match	37.6%	Score 516.8;	DB 13; Length 531;
Best Local Similarity	99.4%;	Prod. No. 1.7e-102;	
Matches 529;	Conservative 0;	Mismatches 2;	Indels 1; Gaps 1;
Db	661	GTGCTGGCCCGGCGCTGCTGGAGAAATTCACGCACTGGGCGCAGTTCCTCGGGCC	720
QY	721	CAGTCCGCTCTGCGCAGAGTGTCTGGAGCTGTGGCCCTGCGGCGAGGTCCTCCCTGGCG	780
Db	61	CAGTCCGCTCTGCGCAGAGTGTCTGGAGCTGTGGCCCTGCGGCGA-66TCTCTCTGGCG	119
QY	781	ATTCGACACCTCCCTGGGCGCAATCTCAAGCGCTGCCCCCTTCTACCTACTTGGACTC	840
Db	120	ATTCGACACCTCCCTGGGCGCAATCTCAAGCGCTGCCCCCTTCTACCTACTTGGACTC	179
QY	841	TTTCAGGTTACTAGAGGCGCTGTCTCCCAACAGTGTCCAGGCTGCGGCTCCCT	900

D	b	180	TTCCAGGATTCACTGAGGGGCCCTCGTCTCCTCCCGCAGATGTCTCCAGGCTTCGGGGCTCCCT	239
O	y	901	GCACAGCTCTGTGGGACCCGGTCCCTCTTGCCCCAACCTCAGCGCTTTGCTCCAGA	960
D	b	240	CGACAGCTCTGTGGGACCCGGTCCCTCTTGCCCCAACCTCAGCGCTTTGCTCCAGA	299
O	y	961	CCTGCCCCCTCCCTAAGAGGCTGCGTGGCCCTGTTACAGTGTTCCTCATCCCACTTAAT	1020
D	b	300	CCTGCCCCCTCCCTAAGAGGCTGCGTGGCCCTGTTACAGTGTTCCTCATCCCACTTAAT	359
O	y	1021	ACAGTATTCACACTTTATCTTACAACACTGCCACCAGCCCACTCTCCAACTCAGCTC	1080
D	b	360	ACAGTATTCACACTTTATCTTACAACACTGCCACCAGCCCACTCTCCAACTCAGCTC	419
O	y	1081	CCCAATCCCCTGACCCCTTTGAGGCCCCCAGATGATCTGACTCCCGCTGGCCACAGACCC	1140
D	b	420	CCCAATCCCCTGACCCCTTTGAGGCCCCCAGATGATCTGACTCCCGCTGGCCACAGACCC	479
O	y	1141	CAGGCACTTGTCTTACTGTACTCTGTGGGCAAGATGGGTCCAGAAAGACC	1192
D	b	480	CAGGCACTTGTCTTACTGTACTCTGTGGGCAAGATGGGTCCAGAAAGACC	531
 RESULT 15 BES77781				
L	OCUS	DEFINITION	BES77781	918 bp mRNA linear EST 12-DEC-2000
A	CESSION	LOCUS	602092080P1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206595 5'	
V	ERSION	KEYWORDS	mRNA sequence.	
K	EYWORDS	SOURCE	BES77781 GI:11651493	
O	RGANISM	house mouse.		
R	EERENCE	Mus musculus		
A	UTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
T	ITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
J	OURNAL	NIH-MGC http://mgc.nci.nih.gov/		
C	OMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: csapbs@mail.nih.gov		
		Tissue Procurement: Jeffrey E. Green, M.D.		
		cDNA Library Preparation: Life Technologies, Inc.		
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNL at:		
		http://image.lnl.gov		
		plate: LLM9767 row: p column: 20		
		High quality sequence start: 17		
		High quality sequence stop: 724.		
F	EATURES	location/Qualifiers		
S	OURCE	1. 918		
		/organism="Mus musculus"		
		/strain="FVB/N"		
		/db_xref="taxon:10090"		
		/clone_image="IMAGE:4206595"		
		/clone_lib="NCI CGAP Co24"		
		/lab_host="DH10B (TI phage-resistant)"		
		/note="Organ: colon; Vector: PCMV-SPORT6; Site_1: NotI;		
		Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.		
		Average insert size 1.6 kb. Constructed by Life		
		Technologies. Note: this is a NCI CGAP Library."		
B	ASE COUNT	ORIGIN		
		153 a 292 c 268 g 204 t 1 others		
Q	UERY MATCH	Best Local Similarity	36.0%; Score 493.6; DB 12; Length 918;	
M	ATCHES	577, Conservative	0; Mismatches 84; Indels 3; Gaps 3;	
G	CTGCGCCCTGCTCGGCTCTCTGCTGCGCGGTGTCAGTTGGGAGACCGGCATGCTCTC	257		
I	GCTGGCCGCTCTGCGCTCTCTGCTGCTGCTGTCAGCTGGGAGCTGGCAGCGCTGTC	60		





PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicopeitliche Y;

XX MPI: 1998-145619/13.

DR P-PSDB; AAM47525.

PT Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts

XX Claim 2; Pages 48-50; 69pp; English.

CC The sequence is that encoding human tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of gene encoding probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.

XX Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;

Query Match 100.0%; Score 1373; DB 19; Length 1373;  
Best Local Similarity 100.0%; Pred. No. 7e-283;  
Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATTTGTAAGCTTTGAATTTCCGCGCGCTCCCGCTCCCGCATCTCTGGG 60  
DB 1 ATGTCATTTGTAAGCTTTGAATTTCCGCGCGCTCCCGCTCCCGCATCTCTGGG 60  
QY 61 TCCCGGATGGGGGGGGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
DB 61 TCCCGGATGGGGGGGGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
QY 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 181 CTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 240  
DB 181 CTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 240  
QY 241 AGCCGGGATGCGTCTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
DB 241 AGCCGGGATGCGTCTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
QY 301 CAGGACCCGTCGAACTGAAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 301 CAGGACCCGTCGAACTGAAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 361 AACCGACTAGTGGGCTCGCAGAGTGCACCTTAAGGCGCGGAGAGAGAGAGAGAG 420  
DB 361 AACCGACTAGTGGGCTCGCAGAGTGCACCTTAAGGCGCGGAGAGAGAGAGAGAG 420  
QY 421 GCGATGCGACCGCACTTATGAAGTTCATCCACGACTGGAAGAGAGAGAGAGAGAG 480  
DB 421 GCGATGCGACCGCACTTATGAAGTTCATCCACGACTGGAAGAGAGAGAGAGAGAG 480  
QY 481 GTGGAGCGGAGAGTGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
DB 481 GTGGAGCGGAGAGTGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 541 TACAACCGCGAGATCGGGGAGATTATATGACACCCGGGCTGGGCTACTACTACTACT 600  
DB 541 TACAACCGCGAGATCGGGGAGATTATATGACACCCGGGCTGGGCTACTACTACTACT 600  
QY 601 CAGGTGCACTTTGATGAGGGGAGAGGCTGTCTACTGTAAGCTGGAAGCTTGTGTGATGT 660  
DB 601 CAGGTGCACTTTGATGAGGGGAGAGGCTGTCTACTGTAAGCTGGAAGCTTGTGTGATGT 660  
QY 661 GTGCTGGGCGCTGGGCTGGGAGAAATTCACGCACTGGGCGAGGAGTCCCTGGGCGC 720  
DB 661 GTGCTGGGCGCTGGGCTGGGAGAAATTCACGCACTGGGCGAGGAGTCCCTGGGCGC 720  
QY 721 CAGCTCGGCTCTGCAAGGTGTGAGGCTGTGGCTGTGGCCCTGGGCGAGGAGTCCCTGG 780  
DB 721 CAGCTCGGCTCTGCAAGGTGTGAGGCTGTGGCTGTGGCCCTGGGCGAGGAGTCCCTGG 780  
QY 781 ATCCGCAACCTCCCTGGGCGCAATCTCAAGGCTCCCTTCTCTACTTCTGGAGCTC 840  
DB 781 ATCCGCAACCTCCCTGGGCGCAATCTCAAGGCTCCCTTCTCTACTTCTGGAGCTC 840  
QY 841 TTCCAGGTTCACTAGAGGGGCGCTGTCCCAAGAGTGTCCAGGCTGGCGGCTCCCT 900  
DB 841 TTCCAGGTTCACTAGAGGGGCGCTGTCCCAAGAGTGTCCAGGCTGGCGGCTCCCT 900  
QY 901 CGACAGCTCTGTGGGCAACCGGCTCCCTGTGCCCCACCTTCAGCGCTTTGTGTCAGA 960  
DB 901 CGACAGCTCTGTGGGCAACCGGCTCCCTGTGCCCCACCTTCAGCGCTTTGTGTCAGA 960  
QY 961 CCGTCCGCTCCCTGTGAGGCTGGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 1020  
DB 961 CCGTCCGCTCCCTGTGAGGCTGGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 1020  
QY 1021 ACAGTATTCGCACTTATATCTTACAACTCCCGCAGGCTCCAGCTCACTACTACTC 1080  
DB 1021 ACAGTATTCGCACTTATATCTTACAACTCCCGCAGGCTCCAGCTCACTACTACTC 1080  
QY 1081 CCCAATCCCTGACCTTTGAGGCGCCAGTGTATCTGATCTCCCTGGGCGACAGACCC 1140  
DB 1081 CCCAATCCCTGACCTTTGAGGCGCCAGTGTATCTGATCTCCCTGGGCGACAGACCC 1140  
QY 1141 CAGGAGTGTGTCTACTGATCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
DB 1141 CAGGAGTGTGTCTACTGATCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
QY 1201 GCACTAAGAGGAGGCTGAGCTGTGGCGAGAGAGGCAAGAGAGGAGGAGGAGGAGGAG 1260  
DB 1201 GCACTAAGAGGAGGCTGAGCTGTGGCGAGAGAGGCAAGAGAGGAGGAGGAGGAGGAG 1260  
QY 1261 TTCCCAATGTGAGGGGCGAGAAACAAGACAGAGCTCCCTTGAAGATTCCTGTGGAT 1320  
DB 1261 TTCCCAATGTGAGGGGCGAGAAACAAGACAGAGCTCCCTTGAAGATTCCTGTGGAT 1320  
QY 1321 TTTTAAACAGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373  
DB 1321 TTTTAAACAGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373

RESULT 2  
ID ABR34881 standard; cDNA: 1364 BP.  
XX ABR34881;  
AC  
XX  
XX  
DT 08-MAY-2002 (first entry)  
XX  
XX Human cDNA encoding secreted protein #19.  
DE  
XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KM coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
KW tissue regeneration; wound healing; burn; haematopoiesis;  
KW myeloid cell deficiency; lymphoid cell deficiency.

XX Homo sapiens.

PN WO200177288-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10224.

XX 06-APR-2000; 2000US-195582P.

XX (GEMV ) GENETICS INST INC.

XX Wong GG, Clark HF, Fiechtel K, Agostino MJ, Howes SH, Resnick RJ,  
PI Gulikota K, Graham JR;

DR WPI: 2002-179321/23.

PT Five hundred and ninety two polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders

PS Claim 1; Page 82; 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins. The polynucleotides can be used as probes for the  
CC identification and isolation of full length cDNA and genomic DNA. The  
CC polynucleotides and proteins can also be used as nutritional supplements.  
CC The proteins are useful in the treatment of various immune deficiencies  
CC and disorders such as viral infections, bacterial infections, fungal  
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
CC and conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment of  
CC burns, incisions and ulcers. The proteins are also useful for regulating  
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
CC Sequences ABK34863-ABK3454 represent polynucleotides of the invention.

XX Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;

Query Match 98.0%; Score 1345.8; DB 24; Length 1364;

Best Local Similarity 99.9%; Pred. No. 4.3e-277;  
Matches 1347; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 TCCGCCCGCGGCTCCCTCCCGATCCCTCGGCTCCCGGATGGGGGCGGTGAGG 84  
DB 5 TCCGCCCGCGGCTCCCTCCCGATCCCTCGGCTCCCGGATGGGGGCGGTGAGG 64  
QY 85 CAGGCACAGCCCGCGCGCCCATGGCGCCGTCGAGCAGAGGGGGGCGGCCCGG 144  
DB 65 CAGGCACAGCCCGCGCGCCCATGGCGCCGTCGAGCAGAGGGGGGCGGCCCGG 124  
QY 145 GGGGAGCCCGGACCGCCTGCTGTCGCTCGCTCGCTGGGCTGGGCTGGGCTGGCC 204  
DB 125 GGGGAGCCCGGACCGCCTGCTGTCGCTCGCTCGCTGGGCTGGGCTGGGCTGGCC 184  
QY 205 TGGCTTGGGCTCCCTGCTGTCGCTGTCGCTGGGAGCGCGGATCGCTGTCGCCAG 264  
DB 185 TGGCTTGGGCTCCCTGCTGTCGCTGTCGCTGGGAGCGCGGATCGCTGTCGCCAG 244  
QY 265 GAGCCGCGCCAGGAGAGTGGTGGCAGAGGAGGAGCAGAGCCCGCGAATGATGCC 324  
DB 245 GAGCCGCGCCAGGAGAGTGGTGGCAGAGGAGGAGCAGAGCCCGCGAATGATGCC 304  
QY 325 CAGACGAGAGAGAGCAGATCTCGGCTTCTGTAACGACTAGTGGGCTCGCAGA 384

DB 305 CAGACGAGAGAGAGCAGATCTCGGCTTCTGTAACGACTAGTGGGCTCGCAGA 364

QY 385 AGTGACCTTAAGGCGCGGAAAACACGGGCTCGAAGAGCATGCGACCCATTATGAGTT 444

DB 365 AGTGACCTTAAGGCGCGGAAAACACGGGCTCGAAGAGCATGCGACCCATTATGAGTT 424

QY 445 CATCCAGACCTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504

DB 425 CATCCAGACCTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484

QY 505 GAGGAGCCAGAAATCAACAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 564

DB 485 GAGGAGCCAGAAATCAACAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 544

QY 565 ATAGTCACCCGGGCTGGGCTTACTACTACTACTACTACTACTACTACTACTACTACT 624

DB 545 ATAGTCACCCGGGCTGGGCTTACTACTACTACTACTACTACTACTACTACTACTACT 604

QY 625 GCTGTACTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 684

DB 605 GCTGTACTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 664

QY 685 GAATTCACGACCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCT 744

DB 665 GAATTCACGACCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCT 724

QY 745 GGGCTGTGGGCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCT 804

DB 725 GGGCTGTGGGCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCT 784

QY 805 CTCAGAGCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGG 864

DB 785 CTCAGAGCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGG 844

QY 865 GTCTCCCAAGAGTCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCC 924

DB 845 GTCTCCCAAGAGTCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCC 904

QY 925 CCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAG 984

DB 905 CCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAG 964

QY 985 CTGGGCTGTACAGTGTTCATCCATCAATCAATCAATCAATCAATCAATCAATCAAT 1044

DB 965 CTGGGCTGTACAGTGTTCATCCATCAATCAATCAATCAATCAATCAATCAATCAAT 1024

QY 1045 AACTCCCGCAGCGCCACTCTCCACTCACTCACTCACTCACTCACTCACTCACTCACT 1104

DB 1025 AACTCCCGCAGCGCCACTCTCCACTCACTCACTCACTCACTCACTCACTCACTCACT 1084

QY 1105 CCCAGTGTGATGATCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCC 1164

DB 1085 CCCAGTGTGATGATCTGCGGCTGAGTCCCTGCGGCTGAGTCCCTGCGGCTGAGTCC 1144

QY 1165 TGTGGGCAAGAGTGGTTCAGAGAGCCCACTTCAAGCACTAAGAGGGGCTGGAGCT 1224

DB 1145 TGTGGGCAAGAGTGGTTCAGAGAGCCCACTTCAAGCACTAAGAGGGGCTGGAGCT 1204

QY 1225 GGCAGAGAGCCAAAGAGTGGGCTTAGGCTCAAGAGTGGGCTCAAGAGTGGGCTGAG 1284

DB 1205 GGCAGAGAGCCAAAGAGTGGGCTTAGGCTCAAGAGTGGGCTCAAGAGTGGGCTGAG 1264

QY 1285 CAAAGCAAGCTCTCCCTTGAAGATTCCTGTGAGATTTTAAACAGATTAATTTTAT 1344

DB 1265 CAAAGCAAGCTCTCCCTTGAAGATTCCTGTGAGATTTTAAACAGATTAATTTTAT 1324

QY 1345 TATTATTGACAAATATTTGATTAATGG 1373

DB 1325 TATTATTGACAAATATTTGATTAATGG 1353

RESULT 3



AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 XX  
 AC AAA49717;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 cDNA clone DNA30879-1152.  
 XX  
 KM PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KM breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KM uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KM central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT /tag= a  
 FT sig\_peptide 58..177  
 FT /tag= b  
 FT mat\_peptide 178..804  
 FT /tag= c  
 XX  
 PN WO20037638-A2.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99WO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 98US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 PA (GENTECH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pitti RM, Wood WI;  
 XX  
 DR WPI: 2000-442668/38.  
 DR P-PSDB; AAY95338.  
 XX  
 PT Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO321, PRO324, PRO328, PRO301, PRO356, PRO362, PRO359 or  
 PT PRO866  
 XX  
 PS Claim 20; Fig 3; 172pp; English.  
 XX  
 CC The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.48). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumor cell to PRO179, PRO207, PRO320, PRO324, PRO356, PRO362, PRO359,  
 CC PRO328, PRO301, PRO326, PRO362, PRO356, PRO359 or PRO866 (see  
 CC AAY95337-49), their agonists or chimERIC polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.  
 XX  
 Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

Query Match 96.2%; Score 1320.2; DB 21; Length 1353;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-271;  
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 49 CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGCGACGACAGCCCCCCCCCATG 108  
 DB 1 CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGCGACGACAGCCCCCCCCCATG 60  
 QY 109 GCCGCCCTCGGAGCCGACGAGCGGAGGGGGCCCGGGGGAGCCGCGACTGCTG 168  
 DB 61 GCCGCCCTCGGAGCCGACGAGCGGAGGGGGCCCGGGGGAGCCGCGACTGCTG 120  
 QY 169 GTCCCGCTCGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGG 228  
 DB 121 GTCCCGCTCGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGG 180  
 QY 229 GTCACTTTGGGAGCGCGGATCGCTGTCGCCGACGAGCCCTGCGAGAGAGCTG 288  
 DB 181 GTCACTTTGGGAGCGCGGATCGCTGTCGCCGACGAGCCCTGCGAGAGAGCTG 240  
 QY 289 GCAGAGGAGGACGAGGACCGCTGCGAGTGAATCCCGACAGAAAGCCAGATCT 348  
 DB 241 GCAGAGGAGGACGAGGACCGCTGCGAGTGAATCCCGACAGAAAGCCAGATCT 300  
 QY 349 GCGCCTTTCTGAACCGACTAGTGGCCCTGCGAGAGTGCACCTAAAGCCGGAACA 408  
 DB 301 GCGCCTTTCTGAACCGACTAGTGGCCCTGCGAGAGTGCACCTAAAGCCGGAACA 360  
 QY 409 CGGGCTCGAAGAGCGATCGACGCCCATATATGAAGTTCATCAAGCTGAGAGAGCA 468  
 DB 361 CGGGCTCGAAGAGCGATCGACGCCCATATATGAAGTTCATCAAGCTGAGAGAGCA 420  
 QY 469 GCGGAGGAGGCTGTGAGCGGAGGAGTGAAGTGTGAGAGAGGAGGAGGAGTCAACAGCT 528  
 DB 421 GCGGAGGAGGCTGTGAGCGGAGGAGTGAAGTGTGAGAGAGGAGGAGGAGTCAACAGCT 480  
 QY 529 ACCCCTCTGCTGATCAACCGCGAGATCGGGAGTTATATGATCAACCGCGCTGCTAC 588  
 DB 481 ACCCCTCTGCTGATCAACCGCGAGATCGGGAGTTATATGATCAACCGCGCTGCTAC 540  
 QY 589 TACCTGTACTGTACAGGTGATGATGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 648  
 DB 541 TACCTGTACTGTACAGGTGATGATGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 649 CTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
 DB 601 CTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 QY 709 TCCCTCGGCGCCGACCTCGCTGCGAGTGTGTGGGCTGTGGGCTGTGGGCTGTGGG 768  
 DB 661 TCCCTCGGCGCCGACCTCGCTGCGAGTGTGTGGGCTGTGGGCTGTGGGCTGTGGG 720  
 QY 769 TCCCTCGGCGCCGACCTCGCTGCGAGTGTGTGGGCTGTGGGCTGTGGGCTGTGGG 828  
 DB 721 TCCCTCGGCGCCGACCTCGCTGCGAGTGTGTGGGCTGTGGGCTGTGGGCTGTGGG 780  
 QY 829 TACTTCGAGCTTTCCAGGTTCAGTGAAGGGGCGCTGCTCCGACAGTGTCCAGGT 888  
 DB 781 TACTTCGAGCTTTCCAGGTTCAGTGAAGGGGCGCTGCTCCGACAGTGTCCAGGT 840  
 QY 889 GCGGCGTCCCTCGAGAGTCTGTGGGACCGCGGCTGCTGCGGCGACCGCTGAGCGGT 948  
 DB 841 GCGGCGTCCCTCGAGAGTCTGTGGGACCGCGGCTGCTGCGGCGACCGCTGAGCGGT 900  
 QY 949 CTTCCTCGAGAGCTGCGCTGCTGTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 1008  
 DB 901 CTTCCTCGAGAGCTGCGCTGCTGTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 960  
 QY 1009 TCCCATATATATACAT 1068  
 DB 961 TCCCATATATATACAT 1020  
 QY 1069 CTTCACTAGTCCCAATCCCTGAGCCCTTTAGAGGCCCGACAGTGAATCTGCCCTG 1128



Db	1021	CGTCTAGTCTCCCAATCCCTGACCTTGTGAGGCCCCAGATGATCTGACCTCCCTG	1080
Qy	1129	GCACGAAACCCCGAGGCGATGTGTCACTGTACTGTACTGTGTGGCAAGATGGGTCCAGAG	1188
Db	1081	GCACGAAACCCCGAGGCGATGTGTCACTGTACTGTGTGGCAAGATGGGTCCAGAG	1140
Qy	1189	ACCCGACTTGAGGCACTAAGAGAGGGGCTGGACCTGGGGGAGAGAACCAAGAGACTGGGC	1248
Db	1141	ACCCGACTTGAGGCACTAAGAGAGGGGCTGGACCTGGGGGAGAGAACCAAGAGACTGGGC	1200
Qy	1249	CTAGGCCAGAGAGTTCGCCAATGTGAGGGGCGAGAAACAAGACMACTCCTCCTTGAGAA	1308
Db	1201	CTAGGCCAGAGAGTTCGCCAATGTGAGGGGCGAGAAACAAGACMACTCCTCCTTGAGAA	1260
Qy	1309	TTCCTGTGATTTTTTAAACAGATATATATTTTATATATATATGTGACAAAATGTTGATA	1368
Db	1261	TTCCTGTGATTTTTTAAACAGATATATATTTTATATATATATGTGACAAAATGTTGATA	1320
Qy	1369	AATGG	1373
Db	1321	AATGG	1325

xx	Thirty five nucleic acids encoding PRO polypeptides, useful for
pt	treating benign or malignant tumours, leukemias and lymphoid
pt	malignancies, inflammatory, angiogenic and immunologic disorders -
xx	
ps	Claim 50; Fig 3; 302pp; English.
xx	
cc	The present invention relates to the isolation of novel human PRO
cc	polypeptides and the polynucleotide sequences encoding them. The
cc	PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
cc	useful for treating benign or malignant tumors (e.g. renal, kidney,
cc	bladder, breast, etc), Leukemias and Lymphoid malignancies, other
cc	disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
cc	macrophagal, stromal and blastocoeal disorders, inflammatory, immune
cc	and angioenic disorders. The polynucleotide sequences are also
cc	useful in gene therapy. ABR40254-ABR40288 encode for the human PRO
cc	polypeptides of the invention.
xx	
s0	Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;
	Query Match            96.2%; Score 1320.2; DB 24; Length 1353;
	Best Local Similarity   99.8%; Pred. No.1,2e-271;
	Matches 1322; Conservative   0; Mismatches   3; Indels   0; Gaps   0
OY	CGATCCCTGCGGGTCCCGGGATGGGGGGCGGTGAAGCAGCAACGCCCCCCTCCCATG 108
Db	1 CGATCCCTGCGGGTCCCCGGATGGGGGGCGGTGAAGCAGCAACGCCCCCCTCCCATG 60
OY	GCGGCGCGTCGGAGCGAAGGGGGGGGGGGGGGGGGGGGACCCGGGCACCGCCCTGCTG 168
Db	61 GCGGCGCGTCGGAGCGAAGGGGGGGGGGGGGGGGGGGGACCCGGGCACCGCCCTGCTG 120
OY	GTCCCGCTGCGCTGGGCTTGCGCTGGCGCTGGCCCTGCTCGGCCCTCTCTGTGGCCGTG 228
Db	121 GTCCCGCTGCGCTGGGCTTGCGCTGGCGCTGGCCCTGCTCGGCCCTCTCTGTGGCCGTG 180
OY	GTCACTTTGGGGAGCGCGGCATCGTTCGCCGACAGAACCTGCCAGAGAAGCTGGTG 288
Db	181 GTCACTTTGGGGAGCGCGGCATCGTTCGCCGACAGAACCTGCCAGAGAAGCTGGTG 240
OY	GCAGAGAGAGACAGACAGCCCTCGGAACAATGCCACACAGAAAGAAAGCCAGATCTCT 348
Db	241 GCAGAGAGAGACAGACAGCCCTCGGAACAATGCCACACAGAAAGAAAGCCAGATCTCT 300
OY	GCGCCCTTCTGACCCGACTAGTTGGGCTCCGACAGATGCACTTAAGCGCGGAAAACA 408
Db	301 GCGCCCTTCTGACCCGACTAGTTGGGCTCCGACAGATGCACTTAAGCGCGGAAAACA 360
OY	GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 468
Db	409 GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 420
OY	GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 420
Db	469 GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 528
OY	GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 528
Db	421 GCGGCTCGAAGAGCGATCGCACCCATTATGAAGTTCAATCCAGACTGGACAGAGCGGA 480
OY	AGCCCTCTGCGCTAACAACCGCCAGATCGGGGAGTTTATAGTACACC GGCGCTCTAC 588
Db	529 AGCCCTCTGCGCTAACAACCGCCAGATCGGGGAGTTTATAGTACACC GGCGCTCTAC 540
OY	AGCCCTCTGCGCTAACAACCGCCAGATCGGGGAGTTTATAGTACACC GGCGCTCTAC 540
Db	481 AGCCCTCTGCGCTAACAACCGCCAGATCGGGGAGTTTATAGTACACC GGCGCTCTAC 540
OY	TACTCTACTGTACAGTGCATTTATATGAGGGGAAGGCTGTACTCTAAGCTGACTTG 648
Db	589 TACTCTACTGTACAGTGCATTTATATGAGGGGAAGGCTGTACTCTAAGCTGACTTG 648
OY	TACTCTACTGTACAGTGCATTTATATGAGGGGAAGGCTGTACTCTAAGCTGACTTG 648
Db	541 TACTCTACTGTACAGTGCATTTATATGAGGGGAAGGCTGTACTCTAAGCTGACTTG 600
OY	CTGGTGTGATGTGTCTGTGCGCTGGCGTCTCTGAGAGAAATTCTACACCACTGCGGCACT 708
Db	649 CTGGTGTGATGTGTCTGTGCGCTGGCGTCTCTGAGAGAAATTCTACACCACTGCGGCACT 708
OY	CTGGTGTGATGTGTCTGTGCGCTGGCGTCTCTGAGAGAAATTCTACACCACTGCGGCACT 708
Db	601 CTGGTGTGATGTGTCTGTGCGCTGGCGTCTCTGAGAGAAATTCTACACCACTGCGGCACT 660
OY	TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 768
Db	709 TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 768
OY	TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 720
Db	661 TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 720
OY	TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 828
Db	769 TCCCTCGGGGAGCGCTGCGCTCTGCGCAGAGTGTCTGGGCTGTGGCCCTGGGCGCAGGG 828



```

Db 721 TCTCTCGTGGATCGGACCGCTCCCTGGGCGCAATCAAGCTGCCCTCCGACAC 780
QY 829 TACTTCGGACTCTTCCAGGTCTACTGAGGGCCCTGCTCTCCACAGTGTCCAGGCT 888
Db 781 TACTTCGGACTCTTCCAGGTCTACTGAGGGCCCTGCTCTCCACAGTGTCCAGGCT 840
QY 889 GCCGGCTCCCTCGACAGTCTCTGGGCGCGGTCCTCCCTCCACAGTGTCCAGGCT 948
Db 841 GCCGGCTCCCTCGACAGTCTCTGGGCGCGGTCCTCCCTCCACAGTGTCCAGGCT 900
QY 949 CTTTGCCTCAGACCTGCCCCCTCTCTAGAGCTGCGCTGGGCTGTTCACGTGTTTCCA 1008
Db 901 CTTTGCCTCAGACCTGCCCCCTCTCTAGAGCTGCGCTGGGCTGTTCACGTGTTTCCA 960
QY 1009 TCCCACTATAATACAGTATATCCACTCTTATTTACACTCCGCCACGCGCCACTCTCCA 1068
Db 961 TCCCACTATAATACAGTATATCCACTCTTATTTACACTCCGCCACGCGCCACTCTCCA 1020
QY 1069 CCTCACTAGCTCCCAATCTCGACCTTTGAGGCGCGCGAGTGTGACTGCCCGCTG 1128
Db 1021 CCTCACTAGCTCCCAATCTCGACCTTTGAGGCGCGCGAGTGTGACTGCCCGCTG 1080
QY 1129 GCCACAGACCCCGGAGGCTGTCTGTCTACTGTACTCTGTGGGCAAGATGGTCCAGAG 1188
Db 1081 GCCACAGACCCCGGAGGCTGTCTGTCTACTGTACTCTGTGGGCAAGATGGTCCAGAG 1140
QY 1189 ACCCACTCTCAGGCACTAAGAGGGGCTGACCTGGGCGGAGAGCCAAAGAGACTGGGC 1248
Db 1141 ACCCACTCTCAGGCACTAAGAGGGGCTGACCTGGGCGGAGAGCCAAAGAGACTGGGC 1200
QY 1249 CTAGCCAGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCCCTTGAGAA 1308
Db 1201 CTAGCCAGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCCCTTGAGAA 1260
QY 1309 TTTCCCTGTGATTTTAAACAGATATATTTATTTATTTATTTATTTGACAAAATGTTGATA 1368
Db 1261 TTTCCCTGTGATTTTAAACAGATATATTTATTTATTTATTTATTTGACAAAATGTTGATA 1320
QY 1369 AATGG 1373
Db 1321 AATGG 1325

```

RESULT 6  
AAV47613  
ID AAV47613 standard; cDNA; 1236 BP.

```

AC AAV47613;
XX
DT 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent gene.
XX
KW ss: TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
KM tissue grafting; vasculatization; apoptosis; autoimmune; birth control.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..750
FT CDS /tag= a
FT /product= "TREPA"
XX
PN W09835061-A2.
XX
PD 13-AUG-1998.
XX
PF 12-FEB-1998; 98WO-US02859.
XX
PR 10-FEB-1998; 98US-0021706.
PR 12-FEB-1997; 97US-0798692.
XX

```

```

PA (ABRO ) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI: 1998-447255/38.
XX
DR P-PSDB: AAM29745.
XX
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
XX treatment of autoimmune disease, tumours and inflammation
XX
PS Claim 11: Page 123-4; 142pp; English.
XX

```

The TNF-related endothelium proliferative agent (TREPA), or its activators or agonists, are used to treat a deficit of TREPA, e.g. to promote wound healing or tissue grafting, by promoting vascularisation, also to induce apoptosis for treating cancer and eliminating autoreactive T cells, as an adjunct to cancer chemotherapy or antiviral treatment. TREPA peptides can also be used to target cytotoxic agents or for affinity isolation of the corresponding receptor, the nucleic acid for which can be used to transform tumour cells to render them more responsive to TREPA and to screen for TREPA mimics.

CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting CC vascularisation), inflammation or a wide range of autoimmune conditions, CC conditions involving abnormal stimulation of epithelial cells (e.g. CC atherosclerosis), for birth control (inhibiting ovulation and placental CC formation) or other angiogenic conditions (e.g. ulcers).

XX

Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Query Match 89.3%; Score 1226.4; DB 19; Length 1236;  
Best Local Similarity 99.5%; Pred. No. 1-Le-251;  
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 106 ATGGCCGCCCGCTGGAGCCAGAGCGGAGGGGCGCCGGGGGAGCCGGGACCGCCCTG 165
Db 1 ATGGCCGCCCGCTGGAGCCAGAGCGGAGGGGCGCCGGGGGAGCCGGGACCGCCCTG 60
QY 166 CTGCTCCGCTGGGCTGGGCTGGGCGCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGG 225
Db 61 CTGCTCCGCTGGGCTGGGCTGGGCGCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGG 120
QY 226 GTGCTCAGTTTGGGAGCGGGGCTGCTGCTCCGCGCAGAGCTGCGCCAGAGAGAGCTG 285
Db 121 GTGCTCAGTTTGGGAGCGGGGCTGCTGCTCCGCGCAGAGCTGCGCCAGAGAGAGCTG 180
QY 286 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 345
Db 181 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
QY 346 CCTGGCCCTTTCTTAACCGAGTATGCGCTCGCAGAGAGTGCACCTTAAGCCCGGAAA 405
Db 241 CCTGGCCCTTTCTTAACCGAGTATGCGCTCGCAGAGAGTGCACCTTAAGCCCGGAAA 300
QY 406 ACAGGGGCTCGAAGAGCGATGCGACGCCATTTGAGTTTCAATCCAGCGAGCTGGACAGAGC 465
Db 301 ACAGGGGCTCGAAGAGCGATGCGACGCCATTTGAGTTTCAATCCAGCGAGCTGGACAGAGC 360
QY 466 GGAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 525
Db 361 GGAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 420
QY 526 TCCAGCCCTCTGCGCTACACCGGAGATGGGGAGTTTATAGTACACCGGGGCTGGCTC 585
Db 421 TCCAGCCCTCTGCGCTACACCGGAGATGGGGAGTTTATAGTACACCGGGGCTGGCTC 480
QY 586 TACTACCTGACTGTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 645
Db 481 TACTACCTGACTGTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 646 TTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 705
Db 541 TTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

```

```

OY 706 AGTTCCTCGGGGCCAGCTCCGCTCTGCGAGGTGTCGGGCTGTTGGCCCTCGGGCA 765
    |||||||
Db 601 AGTTCCTCGGGGCCAGCTCCGCTCTGCGAGGTGTCGGGCTGTTGGCCCTCGGGCA 660
OY 766 GGGTCTCTCCCGGGATCCGACCTCCCTGGGGCCATCTCAAGGCTGCCCTTCTC 825
    |||||||
Db 661 GGGTCTCTCCCGGGATCCGACCTCCCTGGGGCCATCTCAAGGCTGCCCTTCTC 720
OY 826 ACCTACTTGGAGCTCTTCAGAGTTCACAGAGGGGCCCTGGTCTCCCACTGCTCCAG 885
    |||||||
Db 721 ACCTACTTGGAGCTCTTCAGAGTTCACAGAGGGGCCCTGGTCTCCCACTGCTCCAG 780
OY 886 GCTGCGGCTCCCTCTGACAGCTCTGCGGACCCGGTCCCTTGCCTCCCACTCAAGCC 945
    |||||||
Db 781 GCTGCGGCTCCCTCTGACAGCTCTGCGGACCCGGTCCCTTGCCTCCCACTCAAGCC 840
OY 946 GCTCTTGGCTCAGACCTCCGCTCCCTCTGAGAGCTCCGCTGCTTACGTCTTT 1005
    |||||||
Db 841 GCTCTTGGCTCAGACCTCCGCTCCCTCTGAGAGCTCCGCTGCTTACGTCTTT 900
OY 1006 CCATCCACATTAATACAGTATTCCTTATCTTACACTCCCACTCCGCTCACTCT 1065
    |||||||
Db 901 CCATCCACATTAATACAGTATTCCTTATCTTACACTCCCACTCCGCTCACTCT 960
OY 1066 CCACCTCACTAGCTCCCAATCCCTGACCTTTAGAGCCCACTGATCTGACTCCCC 1125
    |||||||
Db 961 CCACCTCACTAGCTCCCAATCCCTGACCTTTAGAGCCCACTGATCTGACTCCCC 1020
OY 1126 CTGCGCAGAGACCCCAAGGCTGTGTCTACTCTGTGGGCAAGATGGGTCCAG 1185
    |||||||
Db 1021 CTGCGCAGAGACCCCAAGGCTGTGTCTACTCTGTGGGCAAGATGGGTCCAG 1080
OY 1186 AAGACCCCACTTCAAGGACTAAGAGGGGCTGACCTGCGGCGAGAAAGCAAGAGACTG 1245
    |||||||
Db 1081 AAGACCCCACTTCAAGGACTAAGAGGGGCTGACCTGCGGCGAGAAAGCAAGAGACTG 1140
OY 1246 GGCCTAGGCGGAGATTCCTCAATGTGAGGGGCGAGAAACAAAGCTCTCCCTTGA 1305
    |||||||
Db 1141 GGCCTAGGCGGAGATTCCTCAATGTGAGGGGCGAGAAACAAAGCTCTCCCTTGA 1200
OY 1306 GAATTCCTGTGATTTTAAACAGATATATTTT 1341
    |||||||
Db 1201 GAATTCCTGTGATTTTAAACAGATATATTTT 1236

RESULT 7
AAD04350
ID AAD04350 standard; cDNA; 1236 BP.
XX
AC AAD04350;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vlnlnerary; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..750
FT CDS /*tag- a
FT /product= "Human TREPA (TNF related endothelium
FT /proliferative agent)"
XX
XX US6207642-B1.
XX
XX 27-MAR-2001.
XX
XX 26-JUN-1998; 98US-0105343.

```

```

XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
PA (ABBO ) ABBOTT LAB.
PI
PI Willey SR;
XX
DR WPI; 2001-280760/29.
DR P-PSDB; AAE00891.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
PS Example 2; Column 73-74; 53pp; English.
XX
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts.
CC The present sequence is a cDNA clone ID #69050 encoding human TREPA.
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Query Match 89.3%; Score 1226.4; DB 22; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1..ie-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 106 ATGCGCCCGCTCGAGACCAAGAGGGGGCGGGGGAGAGCCGGCACCGCCCTG 165
    |||||||
Db 1 ATGCGCCCGCTCGAGACCAAGAGGGGGCGGGGGAGAGCCGGCACCGCCCTG 60
OY 166 CTGCTCCCGCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 225
    |||||||
Db 61 CTGCTCCCGCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 120
OY 226 GTGCTCACTTTGGGAGACCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
    |||||||
Db 121 GTGCTCACTTTGGGAGACCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
OY 286 GTGCGAAGAGAGAGACAGAGACCCGCTCGAAGTGAATCCCAAGAGAAAGCAGAT 345
    |||||||
Db 181 GTGCGAAGAGAGAGACAGAGACCCGCTCGAAGTGAATCCCAAGAGAAAGCAGAT 240
OY 346 CTGCGCCCTTCTGAAACGACTAGTGGCTGCGAGAAAGTGCACCTAAAGCCGGA 405
    |||||||
Db 241 CTGCGCCCTTCTGAAACGACTAGTGGCTGCGAGAAAGTGCACCTAAAGCCGGA 300
OY 406 ACAGGGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
    |||||||
Db 301 ACAGGGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 466 GAGAGCGAGAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
    |||||||
Db 361 GAGAGCGAGAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
OY 526 TCAGAGCCCTTGGGCTCAACACCGCAGATGGGAGATTATAGTCAACCGGCTGG 585
    |||||||
Db 421 TCAGAGCCCTTGGGCTCAACACCGCAGATGGGAGATTATAGTCAACCGGCTGG 480
OY 586 TACTACTGTACTGTCAAGTGGCACTTTGATGAGAGGAGAGAGAGAGAGAGAGAG 645
    |||||||
Db 481 TACTACTGTACTGTCAAGTGGCACTTTGATGAGAGGAGAGAGAGAGAGAGAGAG 540
OY 646 TTGCTGTGATGTGTGTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
    |||||||
Db 541 TTGCTGTGATGTGTGTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

```

```

OY 706 AGTTCCTCGGGGCCCCAGCTCCGCTCTGCGCAGTGTCTGGGCTGTGGCCCTGGGGCCA 765
DB 601 AGTTCCTCGGGGCCCCAGCTCCGCTCTGCGCAGTGTCTGGGCTGTGGCCCTGGGGCCA 660
OY 766 GGGTCTCTCCCTGGGATCCGACACCTCCCTGGGCCCCATCTCAAGCTGCCCCCTTCTCTC 825
DB 661 GGGTCTCTCCCTGGGATCCGACACCTCCCTGGGCCCCATCTCAAGCTGCCCCCTTCTCTC 720
OY 826 ACCTACTCTGGAGCTCTTCCAGGTTACAGAGGGGCCCCCTGGTCTCCCCACAGTGTCCAG 885
DB 721 ACCTACTCTGGAGCTCTTCCAGGTTACAGAGGGGCCCCCTGGTCTCCCCACAGTGTCCAG 780
OY 886 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCGGGTCCCTCTGGCCCACTCCAGCC 945
DB 781 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCGGGTCCCTCTGGCCCACTCCAGCC 840
OY 946 GCTCTTTCCTCGACAGCTGCGCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGTCTT 1005
DB 841 GCTCTTTCCTCGACAGCTGCGCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGTCTT 900
OY 1006 CCATCCCACTAATAATACAGTATCTCCACTCTTATCTTACACTCCCCCACTCCACTCT 1065
DB 901 CCATCCCACTAATAATACAGTATCTCCACTCTTATCTTACACTCCCCCACTCCACTCT 960
OY 1066 CCACCTCACTAGACTCCCAATCCCTGACCTTTGAGGCCCCAGTATCTGACTGCTCC 1125
DB 961 CCACCTCACTAGACTCCCAATCCCTGACCTTTGAGGCCCCAGTATCTGACTGCTCC 1020
OY 1126 CTGGCCACAGACCCCCAGGGGCTTGTCTCACTGCTGTGTGGGCAAGATGGGCTCCAG 1185
DB 1021 CTGGCCACAGACCCCCAGGGGCTTGTCTCACTGCTGTGTGGGCAAGATGGGCTCCAG 1080
OY 1186 AAGACCCCACTTACAGGCTAGAGGGGCTGAGCTGGCGGAGGAGCAAGCAAGAGACTG 1245
DB 1081 AAGACCCCACTTACAGGCTAGAGGGGCTGAGCTGGCGGAGGAGCAAGAGAGACTG 1140
OY 1246 GGCCTAGGCGAGGAGTCTCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTTGA 1305
DB 1141 GGCCTAGGCGAGGAGTCTCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTTGA 1200
OY 1306 GAATTCCTCTGGATTTTAAAAAGATATTATTTT 1341
DB 1201 GAATTCCTCTGGATTTTAAAAAGATATTATTTT 1236

```

RESULT 8  
ID AAX23424 standard; DNA; 1030 BP.  
AAX23424;  
AAX23424;  
18-JUN-1999 (first entry)  
Human TNRL3 DNA.  
Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
developmental abnormality; gestational abnormality; prostate cancer;  
APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
apoptosis; human; ss.  
Homo sapiens.  
Key Location/Qualifiers  
FT 1..627  
FT /\*tag= a  
FT /product= "TNRL3"  
XX MO9911791-A2.  
XX 11-MAR-1999.  
XX 04-SEP-1998; 98WO-0518393.

```

XX 05-SEP-1997; 97US-0924634.
XX (UNIV ) UNIV WASHINGTON.
XX Chaudhary PM;
XX WPI: 1999-205191/17.
XX P-PSDB: AAM93590.
XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX Example VII; Fig 13a. 156pp; English.
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX their active fragments. APO4 is useful for diagnosing prostate cancer
XX by determining levels of APO4 in an individual. Prostate cancer can also
XX be treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selective
XX binding agents, useful in diagnosis/treatment of disease by binding of
XX agents to the polypeptide/active fragment which is extracellular, or
XX expressed on the cell surface. The binding is preferably performed in
XX vivo. APO4 polypeptides/active fragments are also useful for screening
XX for agonists and antagonists by binding and observing the change in APO4
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/active
XX fragments and APO4 signal transducer molecules that specifically interact
XX with a cytoplasmic domain of APO4 and detecting a change in level of APO4
XX activity. The method is performed in vivo or in vitro. APO polypeptides
XX are all useful as immunogens for preparing antibodies. APO4 is also
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell line
XX MCF-7, and induced apoptosis.
XX Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:
XX
XX Query Match 60.7%; Score 833.4; DB 20; Length 1030;
XX Best Local Similarity 99.9%; Pred. No. 4.8e-168;
XX Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 229 CTCAGTTTGGGAGCGGGGACATCGCTGTCGCCGACAGACCTGCCAGAGAGAGCTGGTG 288
DB 1 GTCAGTTTGGGAGCGGGGACATCGCTGTCGCCGACAGAGCTGCCAGAGAGAGCTGGTG 60
OY 289 GCAGAGAGAGACACAGACCCGTCGAGACTGAATCCCGACAGACAGAGAAAGCCAGATCCT 348
DB 61 GCAGAGAGAGACACAGACCCGTCGAGACTGAATCCCGACAGACAGAGAAAGCCAGATCCT 120
OY 349 GCGCCTTTCCTGAAACGACTAGTTGGGCTCGCAGAGTGCACCTTAAAGCCGGAACA 408
DB 121 GCGCCTTTCCTGAAACGACTAGTTGGGCTCGCAGAGTGCACCTTAAAGCCGGAACA 180
OY 409 CGGGCTCGAAGAGCATGGCAGCCCATTTATGAAGTTTCACGACGCTGGACAGACGGA 468
DB 181 CGGGCTCGAAGAGCATGGCAGCCCATTTATGAAGTTTCACGACGCTGGAGAGAGACGGA 240
OY 469 GCGCAGGAGAGGTGTGGAGGCGGACAGTGAATGGCTGGGAGAGAGACAGATCAACAGCTCC 528
DB 241 GCGCAGGAGAGGTGTGGAGGCGGACAGTGAATGGCTGGGAGAGAGACAGATCAACAGCTCC 300
OY 529 AGCCCTCTGCGCTACACCCGCGAGATCGGGAGTTATAGTACACCCGGGCTGGCTTAC 588
DB 301 AGCCCTCTGCGCTACTACCGCAGATCGGGAGTTATAGTACACCCGGGCTGGCTTAC 360
OY 589 TACCTGTACTGTACAGTGCATTTTATGAGAGGGAAGGCTGTCTACTCTAAGCTGACTTG 648
DB 361 TACCTGTACTGTACAGTGCATTTTATGAGAGGGAAGGCTGTCTACTCTAAGCTGACTTG 420
OY 649 CTGCTGATGTGTCTGGCCCTGGCTGCTGAGAGATTTCTAGCCACTGCGGCCAGT 708

```

```

Db 421 CTGGTGAATGATGCTCTGCGCCCTGCGCTGCTGCGAGAAATTTCTGAGCCACTGCGCCACT 480
OY 709 TCCCTGCGGCCCCCAGCTCCGCTCTGCGAGGTGCTGCGCTGTTGGCCCTGCGGCGCAGAG 768
Db 481 TCCCTGCGGCCCCCAGCTCCGCTCTGCGAGGTGCTGCGCTGTTGGCCCTGCGGCGCAGAG 540
OY 769 TCCCTGCGGCCCCCAGCTCCGCTCTGCGAGGTGCTGCGCTGTTGGCCCTGCGGCGCAGAG 828
Db 541 TCCCTGCGGCCCCCAGCTCCGCTCTGCGAGGTGCTGCGCTGTTGGCCCTGCGGCGCAGAG 600
OY 829 TACTTGGAGCTCTTCCAGGTTCTACAGAGGCGCGCTGCTGCTCCAGTCCAGTCCAGGCT 888
Db 601 TACTTGGAGCTCTTCCAGGTTCTACAGAGGCGCGCTGCTGCTCCAGTCCAGTCCAGGCT 660
OY 889 GCGGCTCTCCCTGCGAGAGCTCTGCGGCGAGCCGCTCCCTCTGCGGCGCAGCTCCAGGCT 948
Db 661 GCGGCTCTCCCTGCGAGAGCTCTGCGGCGAGCCGCTCCCTCTGCGGCGCAGCTCCAGGCT 720
OY 949 CTTTCTCCAGAGCTCCGCTCCCTCTGAGAGGCTGCGGCGCTGCTGCTCCAGTCCAGGCT 1008
Db 721 CTTTCTCCAGAGCTCCGCTCCCTCTGAGAGGCTGCGGCGCTGCTGCTCCAGTCCAGGCT 780
OY 1009 TCCCAATTAATATACAGTATTCCTTACAACTCCGCGCCAGCCGCGCACT 1063
Db 781 TCCCAATTAATATACAGTATTCCTTACAACTCCGCGCCAGCCGCGCACT 835

```

## RESULT 9

AAS03964

ID AAS03964 standard; DNA: 898 BP.

AAS03964:

26-SEP-2001 (first entry)

Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA.

TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis; ocular neovascularisation; diabetic retinopathy; neovascular glaucoma; retinoblastoma; retinopathy of prematurity; retrolental fibroplasia; rubeosis; uveitis; macular degeneration; arthritis; Rheumatism; ds; corneal graft neovascularisation; psoriasis; metastatic condition; malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint; KW preneoplastic condition; myocardial angiodysplasia; wound granulation; KW scleroderma; vascular adhesion; telangiectasia; Ischemia; human; KW atherosclerotic plaque neovascularisation; coronary atherosclerosis; peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR; fusion protein.

Homo sapiens.

OS Synthetic.

Location/Qualifiers

Key 52..873

/\*tag- a /product- "Fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and human TWEAK extracellular

domain"

MO200145730-A2.

28-JUN-2001.

19-DEC-2000; 2000MO-US34755.

20-DEC-1999; 99US-0172878.

10-MAY-2000; 2000US-0203347.

(IMMUNEX CORP.

Willey SR.

XX WP1: 2001-417975/44.  
 DR P-PSDB: AAU03499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1: Page 39-40; 46pp; English.

The sequence represents a DNA from the expression vector  
 pDC409-L2-TWEAK, which encodes a fusion protein comprising a growth  
 hormone leader, a leucine zipper multimerisation domain, and the  
 extracellular domain of human TWEAK. The fusion protein was used in  
 the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 from a COS cell human cDNA library. The TWEAK protein is a  
 member of the tumour necrosis factor (TNF) family and induces  
 angiogenesis. TWEAKR may therefore be used to screen for and develop  
 TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 used in the treatment and diagnosis of human disease. The disorders  
 mediated by angiogenesis include ocular disorders characterised by ocular  
 neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 rubeosis, uveitis, macular degeneration and corneal graft  
 neovascularisation, and inflammatory diseases such as arthritis,  
 rheumatism and psoriasis. Other treatable diseases include malignant and  
 metastatic conditions such as sarcomas and carcinomas, benign tumours and  
 preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 scleroderma, vascular adhesions, atherosclerotic plaque  
 neovascularisation, telangiectasia, wound granulation, coronary  
 atherosclerosis, peripheral atherosclerosis and Ischemia.

Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:

Query Match 45.88; Score 629.2; DB 22; Length 898;  
 Best Local Similarity 99.58; Pred. No. 1,4e-124;  
 Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 232 AGTTGGGAGCCGCGGATTCGCTGCGCCAGAGAGCTGCGCCAGAGAGCTGTGGCA 291
Db 250 AGTTGGGAGCCGCGGATTCGCTGCGCCAGAGAGCTGCGCCAGAGAGCTGTGGCA 309
OY 292 GAGGAGGACGAGACCCGCTGCGAATGATCCAGACAGAGAAAGCCAGATCTGCG 351
Db 310 GAGGAGGACGAGACCCGCTGCGAATGATCCAGACAGAGAAAGCCAGATCTGCG 369
OY 352 CCTTCCGAAACGAGTATGTTGCGCTGCGAGAGTGCACCTAAGCCGGAAGACGCG 411
Db 370 CCTTCCGAAACGAGTATGTTGCGCTGCGAGAGTGCACCTAAGCCGGAAGACGCG 429
OY 412 GCTGGAAGAGGATCGAGCCATTTATGAGTTTCATCCAGACCTGAGACGAGGAGCG 471
Db 430 GCTGGAAGAGGATCGAGCCATTTATGAGTTTCATCCAGACCTGAGACGAGGAGCG 489
OY 472 CAGCAGAGTGTGACGCGGAGACAGTGTGCTGCGAGAGAGCCAGATCAACAGCTCAGC 531
Db 490 CAGCAGAGTGTGACGCGGAGACAGTGTGCTGCGAGAGAGCCAGATCAACAGCTCAGC 549
OY 532 CCTGCGGCTACAAACGCGCAGATCGGGAGTTTATATCAACCGGGCTGGCTCTACTAC 591
Db 550 CCTGCGGCTACAAACGCGCAGATCGGGAGTTTATATCAACCGGGCTGGCTCTACTAC 609
OY 592 CTGACTGTGAGGTGCACTTGTGATGAGGAGAGGCTGTACCTGAGAGCTGTGCTG 651
Db 610 CTGACTGTGAGGTGCACTTGTGATGAGGAGAGGCTGTACCTGAGAGCTGTGCTG 669
OY 652 GTGAGTGTGTCTGCGCTGCGCTGCGTGTGAGAGAAATTCACAGCAGTGGCGAGTTC 711
Db 670 GTGAGTGTGTCTGCGCTGCGCTGCGTGTGAGAGAAATTCACAGCAGTGGCGAGTTC 729
OY 712 CTGCGGCGGAGCTCGGCTCTGCGAGGTGTGTGGCTGTGGCCCTGCGGCGAGGCTCC 771

```





```

XX OS Homo sapiens.
XX PM W09514772-A1.
XX PD 01-JUN-1995.
XX PE 11-NOV-1994; 94WO-JP01916.
XX PR 12-NOV-1993; 93JP-0355504.
XX PA (MATSU/) MATSUBARA K.
XX PA (OKUBO/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX DR WPI; 1995-206931/27.
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX PT for diagnosis of abnormal cell function, by preparing cDNA that
XX PT reflects relative abundance of corresp. mRNA in specific human
XX PT tissues
XX PS Claim 1; Page 1067; 2245pp; Japanese.
XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX CC given in AAT19001-T26837 and which is able to hybridise to part of
XX CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX CC sequences were obtained from 3'-directed cDNA libraries prepared
XX CC from various human tissues; synthesis of cDNA was initiated from the
XX CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX CC untranslated sequence is unique to a particular mRNA species, almost
XX CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX CC is constructed so as to reflect accurately the relative abundance of
XX CC different mRNAs in the particular tissue from which it was derived.
XX CC The appearance frequency of a given GS in a cDNA library can be
XX CC determined (esp. using primers and probes derived from the GS
XX CC sequences) as a means of diagnosing abnormal cell function or for
XX CC recognising different cell types.
XX SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;
XX
Query Match 17.9%; Score 245.4; DB 16; Length 282;
Best Local Similarity 97.3%; Pred. No. 5.3e-43;
Matches 257; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
OY 1111 GATCTGACCTCCCTGCGCCACAGACCCCGAGGCGATTGTGTCACCTGCTGTGGG 1170
DB 1 GATCTGACCTCCCTGCGCCACAGACCCCGAGGCGATTGTGTCACCTGCTGTGGG 60
OY 1171 CAAGATGGGTCCAAAGAACCCCACTTCAGGACCTAAGAGGGGCTGGACCTG-GGGCGAG 1229
DB 61 CAAGATGGGTCCAAAGAACCCCACTTCAGGACCTAAGAGGGGCTGGACCTGTGGCGAG 120
OY 1230 GAAGCCAAAGAGACTGGGCTTAGGCGCAGAGATTCACCAATGTAGAGGGGAGAGAAACAAGA 1289
DB 121 GAAGCCAAAGAGACTGGGCTTAGGCGCAGAGATTCACCAATGTAGAGGGGAGAGAAACAAGA 180
OY 1290 CAAGCTCTCCCTTGAGATTCCTGTGAGATTTTAAACAGATATATATTTATATATTA 1349
DB 181 CAAGCTCTCCCTTGAGATTCCTGTGAGATTTTAAACAGATATATATTTTNNATTA 240
OY 1350 TTGTGACAAATGTTGATTAATGG 1373
DB 241 TTGTGACAAATGTTGNTAAATGG 264

```

RESULT 13  
 ABR29540  
 ID ABR29540 standard; cDNA; 195 BP.  
 XX  
 AC ABR29540;

```

XX DT 23-APR-2002 (first entry)
XX DE Colon adenocarcinoma-specific cDNA #66.
XX DE Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX OS Homo sapiens.
XX PM W0200196389-A2.
XX PD 20-DEC-2001.
XX PE 07-JUN-2001; 2001WO-US18574.
XX PR 09-JUN-2000; 2000US-210667P.
XX PR 22-NOV-2000; 2000US-252614P.
XX PA (CORI-) CORIXA CORP.
XX PI Meagher MJ, King GE, Xu J, Secretist H;
XX DR WPI; 2002-098052/13.
XX PT New isolated polynucleotide encoding a polypeptide comprising a portion
XX PT of colon tumour protein, for detection, diagnosis and therapy of human
XX PT colon cancer
XX PS Claim 1; Page 133; 211pp; English.
XX CC The invention relates to an isolated polynucleotide (I) encoding a
XX CC polypeptide (II) comprising a portion of a colon tumour protein. A new
XX CC oligonucleotide (III) that hybridises to (I) is useful for antigen
XX CC determining the presence of a cancer in a patient. (II) or antigen
XX CC presenting cells expressing (I) are useful for stimulating and/or
XX CC expanding T cells specific for a tumour protein, by contacting T cells
XX CC with (I), (II) or antigen-presenting cells that express (I), (II),
XX CC or antigen presenting cells that express (II) are useful for treating
XX CC colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated
XX CC from a patient with (I), (II), or antigen presenting cells that express
XX CC (II), so that T cells proliferate, and administering to the patient an
XX CC effective amount of the proliferated T cells, thus inhibiting the
XX CC development of a cancer in the patient. A new composition is useful for
XX CC stimulating an immune response in a patient. (I) or (II) is useful in
XX CC vaccines and pharmaceutical compositions for prevention and treatment of
XX CC colon cancer and for the diagnosis and monitoring of the cancers. (I),
XX CC (II) or an antibody against (II) is useful for detection, diagnosis and/
XX CC or therapy of human colon cancer. (I) is useful as a probe or primer for
XX CC nucleic acid hybridisation, and in the design and preparation of ribozyme
XX CC molecules for inhibiting expression of (II) in tumour cells. ABR29475-
XX CC ABR29851 represent human colon adenocarcinoma-specific cDNA sequences of
XX CC the invention.
XX SQ Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;
XX
Query Match 14.1%; Score 193.4; DB 24; Length 195;
Best Local Similarity 99.5%; Pred. No. 5.7e-32;
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1130 CCACAGACCCCGAGGCAATGTGTCTACTGTACTGTGTGGCCAGATGGGTCCAGAAGA 1189
DB 1 CCACAGACCCCGAGGCAATGTGTCTACTGTACTGTGTGGCCAGATGGGTCCAGAAGA 60
OY 1190 CCCCACTTCAGGCAATAAGAGGGGCTGACCTGGCGGAGAGACCAAGAGACTGGGCC 1249
DB 61 CCCCACTTCAGGCAATAAGAGGGGCTGACCTGGCGGAGAGACCAAGAGACTGGGCC 120
OY 1250 TAGCCAGAGATTCACCAATGTAGAGGGGCGAGAAACAGACAGAGCTCTCCCTGAGAGAT 1309
DB 121 TAGCCAGAGATTCACCAATGTAGAGGGGCGAGAAACAGACAGAGCTCTCCCTGAGAGAT 180
OY 1310 TCCTGTGATTTT 1324

```



Db 181 TCCCTGTGATTTT 195

RESULT 14

AX53491/c  
ID AX53491 standard; DNA; 114955 BP.

AX53491;

05-JUL-1999 (first entry)

Human adenosine A1 receptor antisense oligonucleotide fragment.

Antisense oligonucleotide; multiple target; antisense treatment;  
impaired respiration; inflammation; lung disease;  
pulmonary vasoconstriction; inflammation; allergic rhinitis;  
acute asthma; allergy; asthma; impaired respiration;  
respiratory distress syndrome; pain; cystic fibrosis;  
pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
colon cancer; breast cancer; lung cancer; pancreatic cancer;  
hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
prostate cancer; ss.

Synthetic.

WO913886-A1.

25-MAR-1999.

17-SEP-1998; 98WO-US19419.

09-JUN-1998; 98US-0093972.

17-SEP-1997; 97US-0059160.

(UYEC-) UNIV EAST CAROLINA.

Myce JM;

WPI; 1999-229400/19.

New antisense oligonucleotides used in treatment of, e.g. pulmonary vasoconstriction

Disclosure; Page 37; 120pp; English.

The specification describes antisense oligonucleotides (AA52869-X55271) directed against at least 2 mRNAs selected from target genes, coding and non-coding regions of RNAs corresponding to target genes, gene initiation codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-end and the juxta-section between coding and non-coding regions and all segments of RNAs encoding proteins associated with one or more diseases, conditions or mixtures. The antisense oligonucleotides may be derived from sequences AA5272-74. These multiple target oligonucleotides (specifically AA5180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, CC hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 5.18; Score 70.4; DB 20; Length 114955;  
Best Local Similarity 32.78; Pred. No. 3.6e-05;  
Matches 182; Conservative 58; Mismatches 316; Indels 0; Gaps 0;

OY 26 CCGCCCGCGGTCCTCCCTCCCGATCTCTGGGTCCCGGATGGGGGCGGTAGAGC 85  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105272 CCGCCCGCGGCGCGCCNNHNNNSCGCCGCGCGCGCCNNHNNNSCGCCGCGC 105213  
OY 86 AGGCACAGCCCCCCCCCATGAGCGCCCGCTCGAGCAGAGAGGGAGGGCGCCG 145  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105212 CCGCGCGCGCGCCNNHNNNSCGCCGCGCGCGCCNNHNNNSCGCCGCGC 105153  
OY 146 GGGAGCGGGGACCGCCCTGTGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCT 205  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105152 GCGCGCGCGCGCCNNHNNNSCGCCGCGCGCGCCNNHNNNSCGCCGCGC 105093  
OY 206 GCTTCGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 265  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105092 GCGCGCGCGCGCCNNHNNNSCGCCGCGCGCGCCNNHNNNSCGCCGCGC 105033  
OY 266 AGCCGCGCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 325  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105032 CCG 104973  
OY 326 AGACAGAGAAAGCCAGGATCTGCGCTTCTGTAACGACTAGTGGCTGCGAGAA 385  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104972 CCGGNNHNNNSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104913  
OY 386 GTGCACCTAAAGCCCGGAAACACBGGCTCGAAGAGCATGCAATGTAAGTTTC 445  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104912 CCG 104853  
OY 446 ATCCAGCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 505  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104852 VGGCCVGGGNNHNNNSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104793  
OY 506 AGAAGCCAGATCAACAGCTCCAGCCCTCTGCGCTCAACAGCGAGATCGGGAGTTA 565  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104792 VGGCCVGGGNNHNNNSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104733  
OY 566 TAGTCACCCCGCGCTGG 581  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104732 NSCGCGCGCGCGCGCG 104717

RESULT 15

AAZ17263  
ID AAZ17263 standard; cDNA; 1337 BP.

AAZ17263;

12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:4735.

Human; gene; gene expression product; diagnosis; therapy; probe;  
detection; mapping; tissue typing; profiling; forensic; cancer;  
genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

Homo sapiens.

WO938972-A2.

05-AUG-1999.

28-JAN-1999; 99WO-US01619.

03-APR-1998; 98US-0080666.

28-JAN-1998; 98US-0072910.

24-FEB-1998; 98US-0075954.

31-MAR-1998; 98US-0080114.

03-APR-1998; 98US-0080515.

(CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

Ctkventakov R, Dickson M, Drmanac R, Drmanac S;



Search completed: May 8, 2003, 07:30:46  
Job time : 346.477 secs

PI Escobedo J, Garcia PD, Garcia V, Gleese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Rita D, Labat I;  
PI Iamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX  
DR WPI: 1999-494092/41.

XX  
PT Novel human genes and their expression products which are  
XX differentially expressed in different cell types

PS Claim 1: Page 2250-2251; 2479pp; English.

XX  
CC The present invention describes a library of human polynucleotides  
CC comprising the sequences given in AA212532 to AA217779. Also described is  
CC a method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.

SO Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

Query Match 4.4%; Score 61; DB 20; Length 1337;

Best Local Similarity 32.1%; Pred. No. 0.0013;

Matches 145; Conservative 0; Mismatches 307; Indels 0; Gaps 0;

OY 690 CTCAGCCACTGGGCGGAGTCCCTCGGGGCCAGCTCGGCTGTGCGAGGTGTGGGCT 749  
DB 382 CANNNNCCNN 441  
OY 750 GTTGGCCCTGGGGGAGGCTCTCGTGGATCCGCAACCTCCCTGGGCCCATCTCAA 809  
DB 442 CCNCCGCCNN 501  
OY 810 GGCTGCCCCCTCTCTACTACTTGGACTTTCAGGTTCTAGAGGGCCCTGCTCTC 869  
DB 502 NNNCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 561  
OY 870 CCCAGAGTGTGTCAGAGGTGCGGCTCCCTCGAGAGCTCTGTGGGCAACCGGCTCC 929  
DB 562 CCCNNCCNN 621  
OY 930 TGCCCCACCTTCAGCGGCTCTTGTCTCCAGAGCTGCGCTCCCTCTAGAGGCTGCTGG 989  
DB 622 CCNCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 681  
OY 990 CCGTGTACAGTGTTCATCCACATATAATAGATATTCACCTTATCTTACACTC 1049  
DB 682 NNNCCNNCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 741  
OY 1050 CCCACAGCGCCACTCTCCACTCTAGCTCCCAATCCCTGAGCCCTTGAGGCCGCCAG 1109  
DB 742 CCNCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 801  
OY 1110 TGATCTGACTTCCCGCTGGCCAGACACCCC 1141  
DB 802 NNNNNNNNNCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 833

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:00 ; Search time 3518.14 Seconds

(without alignments)  
11357.738 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373

Sequence: 1 atgcattgtgatacttga.....gacaaatgtgataatg 1373

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBdb1:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_man:\*

37: em\_htg\_vtl:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320.2	96.2	1353	6 AX201324	AX201324 Sequence
2	1320.2	96.2	1368	9 AF055872	AF055872 Homo sapi
3	1285	93.6	1306	9 AF030099	AF030099 Homo sapi
4	1226.4	89.3	1236	6 AR140407	AR140407 Sequence
5	1096.8	79.9	1651	9 BC019047	BC019047 Homo sapi
6	768.2	56.0	177703	2 AC016876	AC016876 Homo sapi
7	754	54.9	215795	2 AC127470	AC127470 pan trogl
8	629.2	45.8	898	6 AX180714	AX180714 Sequence
9	614.6	44.8	1168	10 AF030100	AF030100 Mus muscu
10	364.2	26.5	15353	2 AC126921	AC126921 Bos tauru
11	309.6	22.5	203083	2 AC069459	AC069459 Mus muscu
12	309.6	22.5	234182	10 AL603707	AL603707 Mouse DNA
13	254.2	18.5	138792	2 AC119115	AC119115 Rattus no
14	246.4	17.9	177555	2 AC130192	AC130192 Sus scrofa
15	242.4	17.7	161428	2 AC126923	AC126923 Canis fam
16	193.4	14.1	195	6 AX379024	AX379024 Sequence
17	158.6	11.6	184026	2 AC098923	AC098923 Rattus no
18	111	8.1	203281	2 AC126237	AC126237 Felis cat
19	100.6	7.3	139405	2 AC126239	AC126239 Felis cat
20	87.2	6.4	7218	6 T6494	Sequence 14
21	77.6	5.7	125020	9 AF429315	AF429315 Homo sapi
22	74.8	5.4	303091	2 AC084799	AC084799 Mus muscu
23	71	5.2	121451	2 AC011931	AC011931 Rattus no
24	70.8	5.2	73948	2 AC022556	AC022556 Homo sapi
25	70.6	5.1	144979	2 AC016280	AC016280 Homo sapi
26	70.4	5.1	1220469	2 AC074307	AC074307 Mus muscu
27	70.2	5.1	1298166	2 AC087563	AC087563 Homo sapi
28	69.2	5.0	.936	8 CNS01AYK	AL113556 Botrytis
29	69.2	5.0	82897	2 AC022595	AC022595 Homo sapi
30	68.4	5.0	202083	2 AC023833	AC023833 Mus muscu
31	68.2	5.0	110737	2 AC011105	AC011105 Homo sapi
32	68.2	5.0	265537	2 AC087228	AC087228 Mus muscu
33	68	5.0	181344	2 AC125880	AC125880 Rattus no
34	68	5.0	300695	2 AC079431	AC079431 Mus muscu
35	67.6	4.9	6624	8 CNS018RP	AL110716 Botrytis
36	67.2	4.9	66624	2 AC101520	AC101520 Mus muscu
37	67.2	4.9	147124	2 AC022768	AC022768 Homo sapi
38	67.2	4.9	177883	2 AC046159	AC046159 Homo sapi
39	67.2	4.9	181988	2 AC090552	AC090552 Homo sapi
40	67.2	4.9	252689	2 AC079433	AC079433 Homo sapi
41	67	4.9	74138	2 AC021272	AC021272 Homo sapi
42	66.8	4.9	62649	2 AC022552	AC022552 Homo sapi
43	66.8	4.9	75628	2 AC021793	AC021793 Homo sapi
44	66.8	4.9	100511	2 AC010774	AC010774 Homo sapi
45	66.8	4.9	205691	2 AC087227	AC087227 Mus muscu

#### ALIGNMENTS

RESULT 1	AX201324	1353 bp	DNA	Linear	PAT 30-AUG-2001
LOCUS	AX201324				
DEFINITION	Sequence 3 from Patent WO0153486.				
ACCESSION	AX201324				
VERSION	AX201324.1	GI:15391154			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.				
AUTHORS	1 (bases 1 to 1353) Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Marsters,S.A., Pan,J., Pittl,R.M., Roy,M.A., Smith,V., Stone,D.M., Watanabe,C.K. and Wood,W.I.				



/product="Apo3/DR3 119and"  
 /protein\_id="AAC39724.1"  
 /db\_xref="GI:3108231"  
 /translation="MAARSRORRRRRRPPGALVPLAIGLALACLLAVSL  
 GSRASLSQERPOEELVABEDDPELAPOTESDPAPEFLRLVPRPSAFKRTK  
 ARRAIAAHVEVHPRPGDQAGVDGTVSGWEAEAINSSPLRYRQIGELVITAGL  
 YLYCQVHEDEKAAVYLKLDLVDVLAIRCIIEESATASASISGLQRLCOVSLAL  
 RPSLSLRITLPMALHAKAPFLTYGLFOVH"  
 BASE COUNT 272 a 443 c 389 g 264 t  
 ORIGIN

Query Match 96.2%; Score 1320.2; DB 9; Length 1368;  
 Best Local Similarity 99.8%; Pred. No. 7.9e-231;  
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 CGATCCCTCGGGTCCGGGATGGGGGGGGGTGAGGAGGACAGAGCCCGCCCGCCGANG 108  
 DB 1 CGATCCCTCGGGTCCGGGATGGGGGGGGGTGAGGAGGACAGAGCCCGCCCGCCGANG 60  
 QY 109 GCGGCCCTGCGAGAGCCAGAGGCGGAGGGGCGCGGGGGGAGCCGGGCAACCGCTGCTG 168  
 DB 61 GCGGCCCTGCGAGAGCCAGAGGCGGAGGGGCGCGGGGGGAGCCGGGCAACCGCTGCTG 120  
 QY 169 GTCCCGCTGCGCTGGGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 228  
 DB 121 GTCCCGCTGCGCTGGGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180  
 QY 229 GTCCAGTTGGGGAGCGGGGATCGCTGCTGCGCCAGAGAGCTGCGCCAGAGAGAGCTG 288  
 DB 181 GTCCAGTTGGGGAGCGGGGATCGCTGCTGCGCCAGAGAGCTGCGCCAGAGAGAGCTG 240  
 QY 289 GCAG 348  
 DB 241 GCAG 300  
 QY 349 GCGCCCTTCTGAG 408  
 DB 301 GCGCCCTTCTGAG 360  
 QY 409 CGGGCTCGAAG 468  
 DB 361 CGGGCTCGAAG 420  
 QY 469 GCGCAG 528  
 DB 421 GCGCAG 480  
 QY 529 AGCCCTCTGCGCTACAG 588  
 DB 481 AGCCCTCTGCGCTACAG 540  
 QY 589 TACCTGACTGACAGTGCAGTTGATGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 648  
 DB 541 TACCTGACTGACAGTGCAGTTGATGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 QY 649 CTGGTGAGAGTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 708  
 DB 601 CTGGTGAGAGTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 660  
 QY 709 TCCCTGAGGAG 768  
 DB 661 TCCCTGAGGAG 720  
 QY 769 TCCCTGAGGAG 828  
 DB 721 TCCCTGAGGAG 780  
 QY 829 TACTTGAGAGTCTCCAGGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888  
 DB 781 TACTTGAGAGTCTCCAGGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
 QY 889 GCGGAGTCCCTGAGAGAGTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948

DB 841 GCGGCTCCCTGCGAGAGAGTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
 QY 949 CTGTGTCAGAGAGTCCCTGCTGAGAGAGTGCCTGGGAGCTGTCACGTGTTTCCA 1008  
 DB 901 CTGTGTCAGAGAGTCCCTGCTGAGAGAGTGCCTGGGAGCTGTCACGTGTTTCCA 960  
 QY 1009 TCCCATATAATACAGTATATCCACATCTTATCTTATACACTCCCGCAGCCACTCTCA 1068  
 DB 961 TCCCATATAATACAGTATATCCACATCTTATCTTATACACTCCCGCAGCCACTCTCA 1020  
 QY 1069 CTTCACTAGTCCCAATCCCTGACCTTTGAGGCCCCAGATATCTGACTCCCTG 1128  
 DB 1021 CTTCACTAGTCCCAATCCCTGACCTTTGAGGCCCCAGATATCTGACTCCCTG 1080  
 QY 1129 GCGCAG 1188  
 DB 1081 GCGCAG 1140  
 QY 1189 ACCCACTTCAGGAG 1248  
 DB 1141 ACCCACTTCAGGAG 1200  
 QY 1249 CTAGGCGAGAGAGTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308  
 DB 1201 CTAGGCGAGAGAGTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
 QY 1309 TCCCTGAGTATTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTT 1368  
 DB 1261 TCCCTGAGTATTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTT 1320  
 QY 1369 AATGG 1373  
 DB 1321 AATGG 1325

RESULT 3  
 AF030099 1306 bp mRNA linear PRI 20-DEC-1997  
 LOCUS  
 DEFINITION  
 ACCESSION AF030099  
 VERSION AF030099.1 GI:2707218  
 KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  
 1 (bases 1 to 1306)  
 Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.M., Scott H.,  
 Hession C., Garcia I., and Browning J.L.

TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)

JOURNAL  
 MEDLINE  
 PUBMED  
 9405449

REFERENCE  
 2 (bases 1 to 1306)  
 Bourdon P., Hession C., Tizard R. and Browning J.  
 Direct Submission  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
 Cambridge, MA 02142, USA

FEATURES  
 source  
 1..1306  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17p13"  
 /tissue\_type="tonsil"  
 18..767  
 /note="ligand in the TNF family; secreted protein; start  
 codon not verified experimentally"  
 /product="TWEAK"  
 /protein\_id="AAC51923.1"  
 /db\_xref="GI:2707219"

## CDS

18..767  
 /note="ligand in the TNF family; secreted protein; start  
 codon not verified experimentally"  
 /product="TWEAK"  
 /protein\_id="AAC51923.1"  
 /db\_xref="GI:2707219"

/translation-"MAARRSORRRRRRGEFETALIVPLAIGLALACGLILAVSL  
 GSRASIASADEPQAELEVAEDQDPSLNPOTSEQDAPLRLVRRRSAPKGRTR  
 ARRAIAAHYVPRPQDQAGQVDTGWSMEERLINSPLRLNRIGEFITRGL  
 YLYVCQHYDEGKAVYKLDLVDGVALCLIEEFESATIASLSICPOLRLCOVSLAL  
 RPSLSLEIRLPLNAHLKAKAPFLYFGLFYH"  
 BASE COUNT 247 a 434 c 368 g 257 t

Query Match 93.6%; Score 1285; DB 9; Length 1306;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-224;  
 Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 CACAGCCCCCCCCCATGGCCCGCTCGAGCCAGAGCCGAGAGGGGCGCCGGGGG 148  
 DB 1 CACAGCCCCCCCCCATGGCCCGCTCGAGCCAGAGCCGAGAGGGGCGCCGGGGG 60  
 QY 149 AGCGGCGACCGCCCTGCTGCTCCGCTGCGCTGGGCTGGGCTGGGCTGGGCTGGC 208  
 DB 61 AGCGGCGACCGCCCTGCTGCTCCGCTGCGCTGGGCTGGGCTGGGCTGGGCTGGC 120  
 QY 209 TCGGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268  
 DB 121 TCGGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 269 CTGGCCAGAGAGAGCTGGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328  
 DB 181 CTGGCCAGAGAGAGCTGGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 QY 329 CAG 388  
 DB 241 CAG 300  
 QY 389 CACCTAAAGGCGGAG 448  
 DB 301 CACCTAAAGGCGGAG 360  
 QY 449 CACGCTGGAG 508  
 DB 361 CACGCTGGAG 420  
 QY 509 AAGCGAATTCACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGTTTATAG 568  
 DB 421 AAGCGAATTCACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGTTTATAG 480  
 QY 569 TCACCCGGGCTGGGCTCTACTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACT 628  
 DB 481 TCACCCGGGCTGGGCTCTACTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACT 540  
 QY 629 TCCTCTGAAGCTGACTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688  
 DB 541 TCCTCTGAAGCTGACTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 QY 689 TCTCAGCAGCTGCGGCGAGTTCTCTGCGGCGCCAGCTCCGCTCTGCGAGTGTCTGGGC 748  
 DB 601 TCTCAGCAGCTGCGGCGAGTTCTCTGCGGCGCCAGCTCCGCTCTGCGAGTGTCTGGGC 660  
 QY 749 TGTGTGGGCTTGGGCGCAGGGTCTCTCTGCGGAGTCCGACCCCTCTGCGGCGCATCTCA 808  
 DB 661 TGTGTGGGCTTGGGCGCAGGGTCTCTCTGCGGAGTCCGACCCCTCTGCGGCGCATCTCA 720  
 QY 809 AGGCTGCGCCCTCTCTACCTACTTGGAGCTCTTCCAGAGTTTCAATGAGGGGCTCTGGTCT 868  
 DB 721 AGGCTGCGCCCTCTCTACCTACTTGGAGCTCTTCCAGAGTTTCAATGAGGGGCTCTGGTCT 780  
 QY 869 CCCCACAGTGTCTCCAGAGCTGCGGCTCTCTGAGAGCTCTTGGGACACCGGCTCTGCTCT 928  
 DB 781 CCCCACAGTGTCTCCAGAGCTGCGGCTCTCTGAGAGCTCTTGGGACACCGGCTCTGCTCT 840  
 QY 929 CTGCGCCACCTCTGAGCGCTTGTCTCAGAGAGCTGCGCTCTCTTGAAGAGCTGCTGCG 988  
 DB 841 CTGCGCCACCTCTGAGCGCTTGTCTCAGAGAGCTGCGCTCTCTTGAAGAGCTGCTGCG 900  
 QY 989 GCGCTGTACAGTGTCTTTCATCCACATTAATACAGTATTCACCTTATCTTACAAT 1048

DB 901 GCTGTTCACGCTGTTTCCATCCACATTAATACAGTATTCCTTATCTTACAAT 960  
 QY 1049 CCCCACAGGCGCAGCTCTCCAGCTGAGTCCGCGCAATCCGTCAGCCCTTGAAGCCCCCA 1108  
 DB 961 CCCCACAGGCGCAGCTCTCCAGCTGAGTCCGCGCAATCCGTCAGCCCTTGAAGCCCCCA 1020  
 QY 1109 GTGATCTGACTCCCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1168  
 DB 1021 GTGATCTGACTCCCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
 QY 1169 GCGAAGATGGGTCCAG 1228  
 DB 1081 GCGAAGATGGGTCCAG 1140  
 QY 1229 GGAAGCCAAAG 1288  
 DB 1141 GGAAGCCAAAG 1200  
 QY 1289 ACAAGCTCTCTCTGAGAAATCCCTGTGATTTTAAACAGATATTTATTTATTTATTT 1348  
 DB 1201 ACAAGCTCTCTCTGAGAAATCCCTGTGATTTTAAACAGATATTTATTTATTTATTT 1260  
 QY 1349 ATTGTGACAAATGTTGATTAATGG 1373  
 DB 1261 ATTGTGACAAATGTTGATTAATGG 1285

RESULT 4  
 ARI14047 1236 bp DNA linear PAT 16-JUN-2001

LOCUS ARI14047  
 DEFINITION Sequence 1 from patent US 6207642.  
 ACCESSION ARI14047  
 VERSION ARI14047.1 GI:14482903  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1236)  
 AUTHORS Wiley,S.R.  
 TITLE Member of the TNF family useful for treatment and diagnosis of disease

JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
 FEATURES Location/Qualifiers  
 SOURCE 1..1236  
 BASE COUNT 225 a 416 c 358 g 237 t

Query Match 89.3%; Score 1226.4; DB 6; Length 1236;  
 Best Local Similarity 99.5%; Pred. No. 1e-213;  
 Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 106 ATGGCCGCCCTCTCGAGCCAG 165  
 DB 1 ATGGCCGCCCTCTCGAGCCAG 60  
 QY 166 CTGCTCCCGCTGCGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 225  
 DB 61 CTGCTCCCGCTGCGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 120  
 QY 226 GTGCTCAGTTTGGGAG 285  
 DB 121 GTGCTCAGTTTGGGAG 180  
 QY 286 GTGCAAG 345  
 DB 181 GTGCAAG 240  
 QY 346 CTGCGGCTTCTCTGAACCGAGTGTGAGGCTGCGAGAGTGAAGTGAAGTGAAGTGAAGTGA 405  
 DB 241 CTGCGGCTTCTCTGAACCGAGTGTGAGGCTGCGAGAGTGAAGTGAAGTGAAGTGAAGTGA 300

OY	406	ACGGGGCTCGAAGAGCGATGGAGCCCATATATGAAGTCAATCCAGACGACCTGGACAGAC	465
Db	301	ACAGGGGCTCGAAGAGCGATGGAGCCCATATATGAAGTCAATCCAGACGACCTGGACAGAC	360
OY	466	GGAGCGCAGGAGGTGTGTGAGCGGAGCACTGTAGTGGCTGGAGGAGACCAATCAACAGC	525
Db	361	GGAGCGCAGGAGGTGTGTGAGCGGAGCACTGTAGTGGCTGGAGGAGACCAATCAACAGC	420
OY	526	TCGAGCCCTGCGCTACACCGCCAGATCGGGAGTTTATATGACACCCGGGCTGGGCTC	585
Db	421	TCGAGCCCTGCGCTACACCGCCAGATCGGGAGTTTATATGACACCCGGGCTGGGCTC	480
OY	586	TACATCCGTACTGTGACAGTGCATCTTATATGAGGGAGGAGCGTCTACTGAACTGAGAC	645
Db	481	TACATCCGTACTGTGACAGTGCATCTTATATGAGGGAGGAGCGTCTACTGAACTGAGAC	540
OY	646	TTGCTGTGTGATGTGTGTCTGGCCCTGGCGCTCGAGAGAAATTCAGACCACTGCGGC	705
Db	541	TTGCTGTGTGATGTGTGTCTGGCCCTGGCGCTCGAGAGAAATTCAGACCACTGCGGC	600
OY	706	AGTTCCCTCGGGGCCCAAGCTCCGCTGCGCAGGTGTCTGGGCTGTGGCCCTGGGGCCA	765
Db	601	AGTTCCCTCGGGGCCCAAGCTCCGCTGCGCAGGTGTCTGGGCTGTGGCCCTGGGGCCA	660
OY	766	GGGTCTCTCCCTGGGATCCGACACCTCCCTGGGGCCATCTAAAGCTGCCCTTCTC	825
Db	661	GGGTCTCTCCCTGGGATCCGACACCTCCCTGGGGCCATCTAAAGCTGCCCTTCTC	720
OY	826	ACCTACTTCGGAGACTTTCAGAGTACTGAGAGGGGCCCTGATCCCAAGTGCATCCAG	885
Db	721	ACCTACTTCGGAGACTTTCAGAGTACTGAGAGGGGCCCTGATCCCAAGTGCATCCAG	780
OY	886	CGTCCGAGCTCCCTCGACAGCTCTGTGGGACCCGGTCCCTCGTCCCGACCCCTCAGCC	945
Db	781	CGTCCGAGCTCCCTCGACAGCTCTGTGGGACCCGGTCCCTCGTCCCGACCCCTCAGCC	840
OY	946	GCTCTTCTCTCAGACGCGCCCTCCCTGAGAGGCTCGCTGGGCTGTTCACAGTGT	1005
Db	841	GCTCTTCTCTCAGACGCGCCCTCCCTGAGAGGCTCGCTGGGCTGTTCACAGTGT	900
OY	1006	CCATCCCAATAAATACAGTATTCACACTTATTTACAACTCCCCACCCGCCCACTCT	1065
Db	901	CCATCCCAATAAATACAGTATTCACACTTATTTACAACTCCCCACCCGCCCACTCT	960
OY	1066	CCACTCTACTAGCTTCCCAATCTCTGACCCCTTGAAGGCCCCAGATGTGACTCCCC	1125
Db	961	CCACTCTACTAGCTTCCCAATCTCTGACCCCTTGAAGGCCCCAGATGTGACTCCCC	1020
OY	1126	CTGGCCACAGACCCCAAGGCGATTGTGTTCACGTACCTGTGGCCAAAGATGGGCTCAG	1185
Db	1021	CTGGCCACAGACCCCAAGGCGATTGTGTTCACGTACCTGTGGCCAAAGATGGGCTCAG	1080
OY	1186	AAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGCGAGGAAGCAAGAAGACTG	1245
Db	1081	AAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGCGAGGAAGCAAGAAGACTG	1140
OY	1246	GACCTTAGGCCAGGAGTCCCAATGTGAGGGGCGAAGAAACAAGACAGCTCTCCCTTGA	1305
Db	1141	GACCTTAGGCCAGGAGTCCCAATGTGAGGGGCGAAGAAACAAGACAGCTCTCCCTTGA	1200
OY	1306	GAATTCCTCTGTGATTTTAAACAGATATTTT 1341	
Db	1201	GAATTCCTCTGTGATTTTAAACAGATATTTT 1236	

RESULT	5
LOCUS	BC019047
DEFINITION	Homo sapiens, Similar to tumor necrosis factor (ligand) superfamily, member 12, clone MGC:20689 IMAGE:4766071, mRNA, complete cds.
ACCESSION	BC019047
VERSION	BC019047.1 GI:17512138
	1651 bp mRNA linear PRI 11-DEC-2001

KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1651)
REFERENCE	Strausberg, R. Direct Submission
AUTHORS	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

Email: [cqabps-remail.nih.gov](mailto:cqabps-remail.nih.gov)  
Tissue Procurement: Louis Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_imgen@nsl.nih.gov](mailto:nisc_imgen@nsl.nih.gov)  
Shevchenko, Y., Weltherby, K. D., Beckstrom-Sternberg, S. M.,  
Benjamin, B., Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S.,  
Dietrich, N. L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, O. L., Mastello, C., Mestrian, S. D., McCloskey, J. C.,  
McGowan, J. E., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P. J.,  
Tjongson, E. E., Touchman, J. W., Tsurgenev, C., Vogt, J. L., Walker, M. A.,  
Zhang, L., H. and Green, E. D.

FEATURES	source	Location/Qualifiers
		1..1651
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="MGC:20669 IMAGE:4766071"
		/tissue.type="Primary B-Cells from Tonsils"
		/clone_idb="NTH MGC_48"
		/lab_host="DH10B-R"
		/note="vector: pOTB7"
CDS		106..510
		/codon_start=1
		/product="Similar to tumor necrosis factor (ligand) superfamily, member 12"
		/protein_id="AAH19047.1"
		/db_xref="GI:17512139"
		/translation="MAARSORRRGRGPEPTALIVPLAGIGIACIGILLAVSISGSRLSISQEPAGQELVAEEODPSSELNPOTEESODPAPFLINRLVPRRASPAPKGRKTRTARRALAHTEVHPHPGGDGAQADSGITCLMP"
BASE COUNT	344 a	517 c 481 g 309 t
ORIGIN		

Query Match	79.9%	Score 1096.8	DB 9	Length 1651
Best Local Similarity	84.8%	Pred. No. 4,1e-190		
Matches 1347	Conservative	0	Mismatches 2	Indels 239
				Gaps 1
0Y	25 TCCGCCGCGGCTCCCCCTCCCCCGATCCTCGGAGTCCCGGAGTGGGGGCGGTGAGG	84		
Db	25 TCCGCCGCGGCTCCCCCTCCCCCGATCCTCGGAGTCCCGGAGTGGGGGCGGTGAGG	84		
0Y	85 CAGGCACAGCCCCCGCCCGCCCATGGCGCGCCCTCGGAGCCAGAGGGGAGGGGGGCGCG	144		
Db	85 CAGGCACAGCCCCCGCCCGCCCATGGCGCGCCCTCGGAGCCAGAGGGGAGGGGGGCGCG	144		
0Y	145 GGGGAGCCGGGACACCGCCCTGCTGGTCCCGCTCGCGCTGGGCTGGGCGCTGGCG	204		
Db	145 GGGGAGCCGGGACACCGCCCTGCTGGTCCCGCTCGCGCTGGGCTGGGCGCTGGCG	204		

205 TGCCCTGGGCTCTGCTGCGCCGCTGCTAGTTTGGGGAGCCGGGATGCTGTCCGCCAG 264  
 205 TGCCCTGGGCTCTGCTGCGCCGCTGCTAGTTTGGGGAGCCGGGATGCTGTCCGCCAG 264  
 265 GAGCCTGCCGAGGAGGAGCTGGTGGCAGAGAGACCGAGACCCGCTGGAACTGAATCCC 324  
 265 GAGCCTGCCGAGGAGGAGCTGGTGGCAGAGAGACCGAGACCCGCTGGAACTGAATCCC 324  
 325 CAGACAG 384  
 325 CAGACAG 384  
 385 AGTGCACCTAAAGCCGAGAAACACGGGCTGGAAGAGCGATGCAAGCCATTATGAAGTT 444  
 385 AGTGCACCTAAAGCCGAGAAACACGGGCTGGAAGAGCGATGCAAGCCATTATGAAGTT 444  
 445 CATCCAGACCTGGACAG 475  
 445 CATCCAGACCTGGACAG 504  
 476 ----- 475  
 505 CCATGATAGTACTAAGTGGTGGAGCCAAAGATTGAACCCAGCTAGAAATGCTGTGACT 564  
 476 ----- 475  
 565 GGACATGCTGTCATGAAGAGGAGAGGCTGCTGTCAGAGGGTGAAGGCTCCATGACAGGG 624  
 476 ----- 475  
 625 CCACATCCAAAAAGGAG 684  
 476 ----- 475  
 685 TTACAGAGAGGCTCAAG 744  
 506 AGGAAAGCAGAAATCAACAGCTCCAGCCCTGCTGCTACACCCGACAGATCGGGAGTTTA 565  
 745 AGGAAAGCAGAAATCAACAGCTCCAGCCCTGCTGCTACACCCGACAGATCGGGAGTTTA 804  
 566 TAGTACACCCGGGCTGGGCTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625  
 805 TAGTACACCCGGGCTGGGCTACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864  
 626 CTGCTACCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 685  
 865 CTGCTACCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924  
 686 AATTCCTAGCAGCTGCGGCGCAGTTCCCTCGGGCCCCAGACTCCGCTGCGCAGGTGTCTG 745  
 925 AATTCCTAGCAGCTGCGGCGCAGTTCCCTCGGGCCCCAGACTCCGCTGCGCAGGTGTCTG 984  
 746 GGTGCTGGGCTGCGGCGCAGTTCCCTCGGGCCCCAGACTCCGCTGCGCAGGTGTCTG 805  
 985 GGTGCTGGGCTGCGGCGCAGTTCCCTCGGGCCCCAGACTCCGCTGCGCAGGTGTCTG 1044  
 806 TCAAGGCTGCCCCCTTCTCTCACTAGTTCGAGCTTCCTGAGGTTGCTGAGGGGCCCTTG 865  
 1045 TCAAGGCTGCCCCCTTCTCTCACTAGTTCGAGCTTCCTGAGGTTGCTGAGGGGCCCTTG 1104  
 866 TCTCCCAACAGTCTGCCAGGCTGCGGCTCCCTTCAGACAGCTCTCTGGGACCCGGTCC 925  
 1105 TCTCCCAACAGTCTGCCAGGCTGCGGCTCCCTTCAGACAGCTCTCTGGGACCCGGTCC 1164  
 926 CCTCTGGCCCAACCTCAAGCCTCTTGTCTGCTCAGACCTGCGCCCTCTGAGAGGCTGCC 985  
 1165 CCTCTGGCCCAACCTCAAGCCTCTTGTCTGCTCAGACCTGCGCCCTCTGAGAGGCTGCC 1224  
 986 TGGGCTGTTTCAAGTGTTCCTCAATCCACATTAATAGATATTCACACTTATCTTACA 1045  
 1225 TGGGCTGTTTCAAGTGTTCCTCAATCCACATTAATAGATATTCACACTTATCTTACA 1284  
 1046 ACTGCCCAACGCGCCACTCTCCACTGACTGACTGCTCCCAATCTCTGACCTTTTGAAGGCC 1105

Db 1285 ACTCCCCACCCGCTCTCTCCACTGACTGCTGCTCCCAATCTCTGACCTTTAGGCCCC 1344  
 1106 CCAGTGAATCTGACTCCGCTGCGCAGAGACCCGAGGCGATGTGTCACTACTCT 1165  
 Db 1345 CCAGTGAATCTGACTCCGCTGCGCAGAGACCCGAGGCGATGTGTCACTACTCT 1404  
 1166 GTGGGCAAGGATGGTCCAGAGAGACCCGAGGCGATGTGTCACTACTCT 1225  
 Db 1405 GTGGGCAAGGATGGTCCAGAGAGACCCGAGGCGATGTGTCACTACTCT 1464  
 1226 GCAGGAAGCCAAAG 1285  
 Db 1465 GCAGGAAGCCAAAG 1524  
 1286 AAGCAGAGCTCTCCCTGAGAGATTCCTGCTGAGATTTTAAACAGATATTTATTTAT 1345  
 Db 1525 AAGCAGAGCTCTCCCTGAGAGATTCCTGCTGAGATTTTAAACAGATATTTATTTAT 1584  
 1346 ATTATTTGACAAATGTTGATTAATG 1373  
 Db 1585 ATTATTTGACAAATGTTGATTAATG 1612  
 RESULT 6  
 AC016876/c 177703 bp DNA linear HTG 06-AUG-2002  
 LOCUS Homo sapiens clone RP11-186B7, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
 DEFINITION unordered pieces.  
 ACCESSION AC016876 GI:21313830  
 VERSION AC016876.5  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULFILL; HTGS\_ACTIVEFIN.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 177703)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldwin, J., Barne, N., Beckerley, R., Boguslavsky, L., Bouhagalter, B.,  
 Brown, A., Castle, A., Collins, S., Collins, S., Collymore, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Domono, M., Donnell, D., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funk, R., Gage, D., Galagan, J.,  
 Galagan, J., Gardyna, S., Grant, G., Hages, B., Heaford, A., Horton, L.,  
 Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C., H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Strange-Rhmann, N., Stojanovic, N., Subramanian, A., Talman, J.,  
 Tesfaye, S., Tirelli, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W.-J., Zimmer, A. and Zody, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 3 (bases 1 to 177703)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
 Barne, N., Bestien, V., Bloom, T., Boguslavsky, L., Bouhagalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S.,  
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,  
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
 McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mienga, Y.,  
 Murphy, T., Naylor, J., Nguyen, C., Nicot, R., Norbu, C., Norman, C., H.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Plunkhang, P., Plier, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,



Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliou, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A., and Zody, M.

# JOURNAL

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 2, 2002 this sequence version replaced g1:15421989.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L3849  
Center clone name: 186\_B\_7

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 48645: contig of 48645 bp in length  
\* 48646 48745: gap of 100 bp  
\* 48746 58277: contig of 9532 bp in length  
\* 58278 58377: gap of 100 bp  
\* 58378 65804: contig of 7427 bp in length  
\* 65805 65904: gap of 100 bp  
\* 65905 79793: contig of 13888 bp in length  
\* 79794 79893: gap of 100 bp  
\* 79894 99493: contig of 19600 bp in length  
\* 99494 99593: gap of 100 bp  
\* 99594 111049: contig of 11456 bp in length  
\* 111050 111149: gap of 100 bp  
\* 111150 125020: contig of 13871 bp in length  
\* 125021 125120: gap of 100 bp  
\* 125121 145109: contig of 19989 bp in length  
\* 145110 145209: gap of 100 bp  
\* 145210 169458: contig of 24249 bp in length  
\* 169459 169558: gap of 100 bp  
\* 169559 177703: contig of 8145 bp in length.

## FEATURES

Location/Qualifiers  
1. 177703  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="RP11-186B7"  
/clone\_lib="RP11-186B7"

BASE COUNT 42862 a 45967 c 46568 g 41254 t 1052 others

## ORIGIN

Query Match 56.0%: Score 768.2; DB 2; Length 177703;  
Best Local Similarity 99.6%: Pred. No. 2.4e-110;  
Matches 770; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 601 CAGGTGACCTTTGATGAGGGAAGGCTCTCACTGAAGCTGAGCTTCTGCTGATGATGAT 660  
DB 37314 CAGGTGACCTTTGATGAGGGAAGGCTCTCACTGAAGCTGAGCTTCTGCTGATGATGAT 37255  
OY 661 GTGCGGCGCTGCTGCTGAGAGAAATTTTCACGCACTGCGGCGCAAGTTCCCTCGGGGCC 720  
DB 37254 GTGCGGCGCTGCTGCTGAGAGAAATTTTCACGCACTGCGGCGCAAGTTCCCTCGGGGCC 37195  
OY 721 CAGGTGACCTTTGATGAGGGAAGGCTCTCACTGAAGCTGAGCTTCTGCTGATGATGAT 780  
DB 37194 CAGGTGACCTTTGATGAGGGAAGGCTCTCACTGAAGCTGAGCTTCTGCTGATGATGAT 37135  
OY 781 ATCCGACCTTCCTGCGGCGCAAGTTTCACGCACTGAGCTGCGGCGCAAGTTTCACGCACT 840

DB 37134 ATCCGACCTTCCTGCGGCGCAAGTTTCACGCACTGAGCTGCGGCGCAAGTTTCACGCACT 37075

OY 841 TTCCAGGTTCACTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

DB 37074 TTCCAGGTTCACTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37015

OY 901 CGACGCTCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

DB 37014 CGACGCTCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36955

OY 961 CTTGCGGCTCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

DB 36954 CTTGCGGCTCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36895

OY 1021 ACAGTATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

DB 36894 ACAGTATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36835

OY 1081 CCCATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

DB 36834 CCCATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36775

OY 1141 CAGGCAATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

DB 36774 CAGGCAATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36715

OY 1201 GCACCTAAGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

DB 36714 GCACCTAAGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36655

OY 1261 TTCCCAATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

DB 36654 TTCCCAATTCCTGCTGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36595

OY 1321 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373

DB 36594 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 36542

RESULT 7  
AC127470 215795 bp DNA linear HTG 17-JUL-2002  
LOCUS Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15  
DEFINITION unordered pieces.  
AC127470.1 GI:2186866  
ACCESSION  
VERSION  
KEYWORDS HTG; PHASE1; HTGS\_DRAFT.  
SOURCE chimpanzee.  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
REFERENCE 1 (bases 1 to 215795)  
AUTHORS Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breese, K., Brinkley, C.,  
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
Haghighi, P., Hansen, N., Ho, S.-T., Iqbal, J.R., Karlins, E., Laric, P.,  
Lee, J.-S., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
Margulies, E.H., Mastaglio, C., Mastaglio, B., Mastaglio, S.D.,  
McCluskey, J.C., McDowell, J., Paguligan, C., Pearson, R.,  
Portnoy, M.E., Prasad, A., Schneider, M.G., Stantibop, S., Thomas, J.W.,  
Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A.,  
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 215795)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
DIRECT SUBMISSION  
SUBMITTED (17-JUL-2002) NIH Intramural Sequencing Center, 8717  
Grovermont Circle, Gaithersburg, MD 20877, USA  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC

```

Web site: http://www.niscc.nih.gov
Contact: niscc@zooeinhri.nih.gov
-----
Project Information
Center project name: cms
Center clone name: 145D13
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214353; sum-of-contigs
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

Db 200176 TTTTAAACAGATATTATTTATATGTGACAAAGTGTGATAATGG 200229

RESULT 8  
 AX180714  
 LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001  
 DEFINITION Sequence 1 from Patent WO0145730.  
 ACCESSION AX180714  
 VERSION AX180714.1 GI:15132570  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1 (bases 1 to 898)  
 AUTHORS Wiley,S.R.  
 TITLE Tweak receptor  
 JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;  
 IMMUNEX CORPORATION (US)  
 FEATURES  
 Location/Qualifiers  
 source 1..898  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 CDS 52..873  
 /note="human TWEAK fusion protein construct"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC50004.1"  
 /db\_xref="GI:15132571"  
 /translation="MATSRTSLIAFLGLPWLQESANTSSDMKOIEDKIEHLS  
 KIVHENEIARIKLIIGERTRSLSRASLSAOEPAOEIYAEODSELPOTRES  
 ODPAPFLNRYPRRSAPGRKTRARAIHAHYEYHPRGDOAGADGTVSGWEA  
 RINSSSPILRYNIOGEFLTVIRAGLYLYICOVHEDSGKAVYLIKLDLYGVATRLTEE  
 FSATASPLIGPOLRICOVSGILAIRPGSLIRKTRIPMAHAKRAPIYTFGLFOVH"

BASE COUNT 187 a 266 c 267 g 178 t  
 ORIGIN

Query Match 45.8%; Score 629.2; DB 6; Length 898;  
 Best Local Similarity 99.5%; Pred. No. 7.6e-105;  
 Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 232 AGTTGGGAGCCGGGCGCTGTCGCCAGAGAGCCGCGGAGAGAGAGCTGTGGCA 291  
 |||||  
 Db 250 AGTTGGGAGCCGGGCGCTGTCGCCAGAGAGCCGCGGAGAGAGAGCTGTGGCA 309  
 |||||  
 QY 292 GAGGAGACACGAGCCGCGGAGCTGAATCCCAACAGAGAGAGAGAGCTGTGGC 351  
 |||||  
 Db 310 GAGGAGACACGAGCCGCGGAGCTGAATCCCAACAGAGAGAGAGAGCTGTGGC 369  
 |||||  
 QY 352 CCTTCCTGAACGACTAATTCGGGCTGCGAGAGTGCACCTAAAGCCGGAACACAGG 411  
 |||||  
 Db 370 CCTTCCTGAACGACTAATTCGGGCTGCGAGAGTGCACCTAAAGCCGGAACACAGG 429  
 |||||  
 QY 412 GCTCGAAGAGCATGCGACCCCATTAATGATGATCCAGACTGAGACAGAGAGG 471  
 |||||  
 Db 430 GCTCGAAGAGCATGCGACCCCATTAATGATGATCCAGACTGAGACAGAGAGG 489  
 |||||  
 QY 472 CAGGACAGTGTGAGAGGAGCACTGAGTGGCTGGAGAGAGAGAGAGTCAACAGCTCCAGC 531  
 |||||  
 Db 490 CAGGACAGTGTGAGAGGAGCACTGAGTGGCTGGAGAGAGAGAGTCAACAGCTCCAGC 549  
 |||||  
 QY 532 CCTTCGCGCTACACCGCCAGATCGGGAGTTTATGATACCGCGGCTGGGCTCTACTAC 591  
 |||||  
 Db 550 CCTTCGCGCTACACCGCCAGATCGGGAGTTTATGATACCGCGGCTGGGCTCTACTAC 609  
 |||||  
 QY 592 CTGTACTGTCAAGTGCATTTGATGAGGAGAGAGCTGTCTACTCTGAAGCTGAGCTGTG 651  
 |||||  
 Db 610 CTGTACTGTCAAGTGCATTTGATGAGGAGAGAGCTGTCTACTCTGAAGCTGAGCTGTG 669  
 |||||  
 QY 652 GTGAGATGTGTGCTGGCCCTGCGCTGCTGAGAGAGATTTCTACGCCACTGCGCCAGTTCC 711  
 |||||  
 Db 670 GTGAGATGTGTGCTGGCCCTGCGCTGCTGAGAGAGATTTCTACGCCACTGCGCCAGTTCC 729

QY 712 CTCGGGCCCCAGCTCCGCTGTCGCCAGAGTGTGGGCTGTGGCCCTGCGGACAGGTC 771  
 |||||  
 Db 730 CTCGGGCCCCAGCTCCGCTGTCGCCAGAGTGTGGGCTGTGGCCCTGCGGACAGGTC 789  
 |||||  
 QY 772 TCCCTGCGGATCCGACACCCCTCCCTGGGCGCATCTCAAGCTGCCCCCTTCTCACTAC 831  
 |||||  
 Db 790 TCCCTGCGGATCCGACACCCCTCCCTGGGCGCATCTCAAGCTGCCCCCTTCTCACTAC 849  
 |||||  
 QY 832 TTCGAGCTCTTCAGAGTTTACTGAGAGGCGCCCTGG 865  
 |||||  
 Db 850 TTCGAGCTCTTCAGAGTTTACTGAGAGGCGCCGCGG 883  
 |||||

RESULT 9  
 AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS AF030100  
 DEFINITION Mus musculus TWEAK mRNA, partial cds.  
 ACCESSION AF030100  
 VERSION AF030100.1 GI:2707220  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1168)  
 Hession,C., Garcia,I. and Browning,J.L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32410-32410 (1997)  
 9405449  
 PUBMED  
 JOURNAL  
 MEDLINE  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN

Query Match 44.8%; Score 614.6; DB 10; Length 1168;  
 Best Local Similarity 75.8%; Pred. No. 3.4e-102;  
 Matches 906; Conservative 0; Mismatches 219; Indels 70; Gaps 9;

QY 178 GGGCTGGGCTGGGCGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 237  
 |||||  
 Db 2 GGGCTGGGCTGGGCGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 61  
 |||||  
 QY 238 GGGAGCCGCGCATGCTGTGCTCCGCGAGAGCTGCGCCAGAGAGAGCTGTGCGCAGAGAG 297  
 |||||  
 Db 62 GGGAGCTGGGCAACGCTGTCTGCCAGAGAGCTTCTCAGAGAGAGCTGTGACACAGAGAG 121  
 |||||  
 QY 298 GACCAAGACCCGCTCGGAAGTGAATCCCAAGAGAGAGAGAGAGAGAGATCTGCGCTTTC 357  
 |||||  
 Db 122 CGCGGGAGAGCCCTCGAAGTGAATCCCAAGAGAGAGAGAGAGAGATCTGCTTTC 181  
 |||||  
 QY 358 CTGAACCGACTGTGCTGCGGCTGCGAGAGTGTGACTTAAGGCCGGAACACGGGCTGCA 417

Db 182 TTGAAACAACTAGTCGGCCCTCGAAGAGTCTCTTAAGCCCGAAGGCCGCGCTTCGC 241  
 QY 418 AGACGATCGAGCCCATTTATGAAATTCACAGACCTGGACAGGACGGACGCA 477  
 Db 242 CGACTATTCAGACCCCATTTATGAGTTTCATCTCGCCAGACAGATGAGACACACCA 301  
 QY 478 GGTGTGACGGGACAGTGTGCTGGAGAGAACCCAAATCAACAGCTCCAGCCCTTG 537  
 Db 302 GGTGTGATGGAGACAGTGTGCTGGAGAGAACCAAAATCAACAGCTCCAGCCCTTG 361  
 QY 538 CGGTAAACCCGCAATCGGGAGATTTTACTACCCGGCTGGGCTTACTACTGTAC 597  
 Db 362 CGGTAAACCCGCAATCGGGAGATTTTACTACATCAAGGCTGGGCTTACTACTGTAC 421  
 QY 598 TGTGAGTGCATTTGATGAGGAGAGAGCTCTACCTGAGAGTGCATCTCTGATGAT 657  
 Db 422 TGTGAGTGCATTTGATGAGGAGAGAGCTCTACCTGAGAGTGCATCTCTGATGAT 481  
 QY 658 GGTGTGCTGGCCCTGGCTGCTGGAGAGAAATTCACAGCTCCAGCCCTTG 717  
 Db 482 GGTGTGCTGGCCCTGGCTGCTGGAGAAATTCACAGCTCCAGCCCTTG 541  
 QY 718 CCCAGCTCGGCTGCTGGAGAGTGTGCTGGAGAGCTGGGCTGGGCTTCCTCG 777  
 Db 542 CCCAGCTCGGCTGCTGGAGAGTGTGCTGGAGAGCTGGGCTGGGCTTCCTCG 601  
 QY 778 CGGATCCGACCCCTCGGCTGGAGAGTGTGCTGGAGAGCTGGGCTGGGCTTCCTCG 837  
 Db 602 CGGATCCGACCCCTCGGCTGGAGAGTGTGCTGGAGAGCTGGGCTGGGCTTCCTCG 661  
 QY 838 CTCTTCAGAGTTCACTGAGGGGCTGCTGCTCCACAGTGTCCAGGCTGGGCTGC 897  
 Db 662 CTCTTCAGAGTTCACTGAGGGGCTGCTGCTCCACAGTGTCCAGGCTGGGCTGC 721  
 QY 898 -----CCGACGACCTGCTGGGACCCGCTGCTGGGACCCGCTGCTGGGCTGCT 952  
 Db 722 AGGACATACACACCTGCTGGGACCCGCTGCTGGGACCCGCTGCTGGGCTGCT 780  
 QY 953 GCTCCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012  
 Db 781 GGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837  
 QY 1013 ACATTAATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072  
 Db 838 -----ACAGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891  
 QY 1073 ACTAGCTCCCATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1132  
 Db 892 ACTAGCTCCCATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 936  
 QY 1133 CAGACCCCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192  
 Db 937 CACCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968  
 QY 1193 CACTTCAGGCTGAGAGGGGCTGAGACCTGGGCTGAGAGAGGATGGGCTGAG 1252  
 Db 969 -----CAGGACGATGAGAGGGGCTGAGACCTGGGCTGAGAGAGGATGGGCTGAG 1023  
 QY 1253 GCCAGAGATTCGCAATGATGAGGGGCGAGA-AACAAGACAGCTGCTGCTGAGAAATTC 1311  
 Db 1024 GCCAGAGATTCGCAATGATGAGGGGCGAGAAGAGCTGCTGCTGAGAAATTC 1079  
 QY 1312 CCGTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1366  
 Db 1080 CCGTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1132  
 RESULT 10  
 AC126921 153553 bp DNA linear HTG 10-JUL-2002  
 LOCUS Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered  
 DEFINITION pieces.

ACCESSION AC126921  
 VERSION AC126921.1 GI:21724098  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE COW,  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 153553)  
 AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,  
 Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
 Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P.,  
 Lee-Ilin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
 Margulies, E.H., Mastello, C., Masker, B., Mastrian, S.D.,  
 McElroy, J.C., McDowell, J., Paquirigan, C., Pearson, R.,  
 Portnoy, M.E., Prasad, A., Scheuler, M.G., Stantirpop, S., Thomas, J.W.,  
 Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Wetherby, K.D., Wiggin, L., Young, A., Zhang, L., H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 153553)  
 TITLE Direct Submission  
 JOURNAL Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
 AUTHORS Grovumont Circle, Galtersburg, MD 20877, USA  
 JOURNAL Genome Center  
 COMMENT Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc-zoo@nigrl.nih.gov  
 Project Information  
 Center project name: ddi  
 Center clone name: 045D24  
 ----- Summary Statistics -----  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 146066 bases at least Q40  
 Consensus quality: 147748 bases at least Q30  
 Consensus quality: 148824 bases at least Q20  
 Insert size: 151000; agarose-fp  
 Insert size: 152453; sum-of-ctrls  
 Quality coverage: 8.80x in Q20 bases; agarose-fp  
 Quality coverage: 8.72x in Q20 bases; sum-of-ctrls  
 ----- NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

Contig	Length	Quality
1	2312	contig of 2312 bp in length
2	2313	gap of unknown length
3	2413	gap of unknown length
4	5841	contig of 3429 bp in length
5	5842	gap of unknown length
6	5942	contig of 2494 bp in length
7	8436	gap of unknown length
8	8536	contig of 7264 bp in length
9	15799	gap of unknown length
10	15800	contig of 9325 bp in length
11	15900	gap of unknown length
12	25224	contig of 7180 bp in length
13	25325	gap of unknown length
14	32504	contig of 17079 bp in length
15	32605	gap of unknown length
16	32605	contig of 8366 bp in length
17	40971	gap of unknown length
18	41070	contig of 15520 bp in length
19	41071	gap of unknown length
20	56591	gap of unknown length
21	56591	contig of 17079 bp in length
22	73770	gap of unknown length
23	73770	contig of 16990 bp in length
24	90859	contig of 16990 bp in length



Center clone name: RP23-168PS  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821  
 Chemistry: Dye-terminator Big Dye; 48% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 212648 bases at least Q40  
 Consensus quality: 218902 bases at least Q30  
 Consensus quality: 222384 bases at least Q20  
 Estimated insert size: 210556; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 7 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 62152: contig of 62152 bp in length  
 \* 62153 62252: gap of unknown length  
 \* 62253 118772: contig of 56520 bp in length  
 \* 118773 118872: gap of unknown length  
 \* 118873 148924: contig of 30052 bp in length  
 \* 148925 149024: gap of unknown length  
 \* 149025 167231: contig of 18207 bp in length  
 \* 167232 167331: gap of unknown length  
 \* 167332 189907: contig of 22576 bp in length  
 \* 189908 190007: gap of unknown length  
 \* 190008 196537: contig of 6530 bp in length  
 \* 196538 196637: gap of unknown length  
 \* 196638 203083: contig of 6446 bp in length.

FEATURES  
 source  
 1. 203083  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-168PS"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others  
 ORIGIN

Query Match 22.5%; Score 309.6; DB 2; Length 203083;  
 Best Local Similarity 72.0%; Pred. No. 9.5e-47;  
 Matches 556; Conservative 0; Mismatches 149; Indels 67; Gaps 9;

QY 601 CAGGTGACCTTGATGAGGGGAGGAGGCTCTCACTGAAAGCTGGAAGCTGCTGTGATGCT 660  
 |||||  
 Db 43019 CAGGTGACCTTGATGAGGGGAGGAGGCTCTCACTGAAAGCTGGAAGCTGCTGTGATGCT 42360  
 QY 661 GTGTGGCCCTGCGCTGCTGAGGAATTCACGCACTGCGGCGAGTTCCCTGGGGCC 720  
 |||||  
 Db 42959 GTGTGGCCCTGCGCTGCTGAGGAATTCACGCACTGCGGCGAGTTCCCTGGGGCC 42300  
 QY 721 CAGTCCGCTCTGCAAGGTCTGTGGGCTGTGGCCCTGCGGCGAGGCTCTCCCTGGCG 780  
 |||||  
 Db 42899 CAGTCCGCTCTGCAAGGTCTGTGGGCTGTGGCCCTGCGGCGAGGCTCTCCCTGGCG 42840  
 QY 781 ATCCGCACTCTCCCTGGGCGCATCTCAAGCTGCGGCGCTTCTCCACCTACTTGGAGTC 840  
 |||||  
 Db 42839 ATCCGCACTCTCCCTGGGCGCATCTCAAGCTGCGGCGCTTCTCCACCTACTTGGAGTC 42780  
 QY 841 TTCAGGTCACAGAGGGGCTGTGTCCTCCCAACAGTCGTCAGGCTGCGGCTCC--- 897  
 |||||  
 Db 42779 TTCAGGTCACAGAGGGGCTGTGTCCTCCCAACAGTCGTCAGGCTGCGGCTCC--- 42720  
 QY 898 -CCTGACACAGCTCTGTGGGCGAGGCTCCCTGCGGCGCACTCAAGCGCTTTTGGCT 955  
 |||||  
 Db 42719 AGCATGACACAGCTCTGTGGGCGAGGCTCCCTGCGGCGCACTCAAGCGCTTTTGGCT 42661

QY 956 CCAGACCTGCCCCCTCCCTCTGAGAGGCTGCGGCGCTGCTACGCTTTTCCATCCACGA 1015  
 |||||  
 Db 42660 CCAGACCTGCCCCCTCCCTCTGAGAGGCTGCGGCGCTGCTACGCTTTTCCATCCACGA 42602  
 QY 1016 TAAATACAGTATTCCTCCATCTTATCTTACACCTCCCGAGCCGACCTCCACTACT 1075  
 |||||  
 Db 42601 GA-----CGTATCCCTGCTCTTCTTTTAACTATCCATCCACACATATCCACTCACT 42547  
 QY 1076 AGCTCCCAACATCCCTGACCTTTGAGGCGCCCACTGATCTGACCTCCCGCTGGCCACG 1135  
 |||||  
 Db 42546 AGCTCCCAACATCCCTGACCTTTGAGGCGCCCACTGATCTGACCTCCCGCTGGCCACG 42502  
 QY 1136 ACCCCAGAGGCTATGTTGTTACTCTGCTGTGGGCAAGATGGGTCCAGAACCCAC 1195  
 |||||  
 Db 42501 CCGACACAGTATTTATGCTTTGTCAC----- 42473  
 QY 1196 TTCAGGCTCAAGAGGGGCTGAGCTGCGGCGAGGACCAAGACCTGGGCTAGGCGC 1255  
 |||||  
 Db 42472 -CAGGCTCAAGATGGGCTGAGCTGCGGCGAGGACCAAGACCTGGGCTAGGCGC 42415  
 QY 1256 AGAGTTCCCAATGTGAGGGGCGAGGAG-MACAAGACAGCTCTCCCTTGAATTCCT 1314  
 |||||  
 Db 42414 AGAGTTCCCAATGTGAGGGGCGAGGAG-MACAAGACAGCTCTCCCTTGAATTCCT 42359  
 QY 1315 GTGATTTTAAACAGATATTTATTTATTTATTTGTCACAAATGTTGA 1366  
 |||||  
 Db 42358 GTGATTTTGA--AGATACTATTTATTTATTTATTTGTCACAAATGTTGA 42309

RESULT 12  
 AL603707/c 234182 bp DNA linear ROD 17-NOV-2001  
 LOCUS  
 DEFINITION  
 Mouse DNA sequence from clone RP23-422L16 on chromosome 11,  
 complete sequence.  
 AL603707  
 AL603707.5 GI:17017790  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On Nov 20, 2001 this sequence version replaced GI:16605765.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; SW:  
 SWISSPROT; Tr: TrEMBL; Wp: WormPeP; Information on the WormPeP  
 database can be found at  
<http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-422L16 is  
 from the RPI-23 Mouse PAC Library  
 constructed by the group of Plietier de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6

FEATURES  
 source  
 1. 234182  
 Location/Qualifiers

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_11b="RPC1-23"
complement(84050..84131)
/feature="Sequence from uni-directional primer reads and dgrp
big dye terminator reads only."
misc_feature
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

```

```

Query Match      22.5%; Score 309.6; DB 10; Length 234182;
Best Local Similarity 72.0%; Pred. No. 9,4e-47;
Matches 556; Conservative 0; Mismatches 149; Indels 67; Gaps 9;

OY 601 CAGGTGACCTTTGATGAGGGGAGAGGCTGTCTACCTGAAAGCTGACTTGTGTGATGCT 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66323 CAGGTGACCTTTGATGAGGGGAGAGGCTGTCTACCTGAAAGCTGACTTGTGTGATGCT 66264

OY 661 GATGAGGCGCCCTGCGCTGCTGAGGAATTCAGGCACTGCGGCGGCGGAGTCCCTCGGCGCC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66263 GATGAGGCGCCCTGCGCTGCTGAGGAATTCAGGCACTGCGGCGGCGGAGTCCCTCGGCGCC 66204

OY 721 CAGCTCCGCTCTGCGAGGTGTCTGTGGCTGTGGCCCTGCGGCGAGGGTCTCTCCCTGCGG 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66203 CAGCTCCGCTCTGCGAGGTGTCTGTGGCTGTGGCCCTGCGGCGAGGGTCTCTCCCTGCGG 66144

OY 781 ATCCGACACCTTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCCTCACTTCTGAGACTC 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66143 ATCCGACACCTTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCCTCACTTCTGAGACTC 66084

OY 841 TTCGAGGTTCACCTAGAGGGGCGCTGCTGCTCCACAGTGTCTCCAGGCTCGCGGCTCC--- 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66083 TTCGAGGTTCACCTAGAGGGGCGCTGCTGCTCCACAGTGTCTCCAGGCTCGCGGCTCC--- 66024

OY 898 ---CCTGACAGCTCTCTGTGGGACCGCGTCCCTGCGCCACCTCAGCGCTCTTGTCT 955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66023 ACCTACACACACACTCCCTACCCCAACCCCACTCCGACCCCTC---GCTGCTCTGCT 65965

OY 956 CCAGACCTGCGCCCTCCCTCTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65964 CCAGCTCTGCTCTCTCTCTC---AAAGGACGACGAGCTGTTCACAGTGTTCCTATTCACA 65906

OY 1016 TAAATACAGTATTCACCTATCTATCACTACCTCCCGCCAGCTCTCCACCTCACT 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65905 GA-----GATATCTTGTCTCTTCTTAACTCCATCCACACACACTATCCACCTCACT 65851

OY 1076 AGCTCCCAATCCCTGACCTCTTGAAGGCGCCAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65850 AGCTCCCAATCCCTGACCTCTTGAAGGCGCCAGTGTCTGCTGCTGCTGCTGCTGCTGCT 65806

OY 1136 ACCCCAGGCGCATGTGTCTCACTGTACTGTGTGGGACAGAGTGGTCCAGAAAGACCCAC 1195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65805 CCGACCGAGCTGTGTCTTGACTTGTGAC----- 65777

OY 1196 TTCAGGCACTAAGAGGGGCTGAGCTGCGGCGGAGGAGCAAGAGAGTGGGCTAGGCGC 1255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65776 ---CAGGCACTAGATGGGCTGAGCTGTGTGGGAGGAGGACAGAAAGTGGGAGTGGGCGC 65719

OY 1256 AGGAGTCCCAATGTGAGGGGCGAGA--AACAGCAAGAGCTCTCTTCTGAGAAATTCCT 1314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65718 AGAAGTCCCAACTGTGAGGGGAGAGAGCTGGGAGCAAGCTCTCTCTGGA-----TCCCT 65663

OY 1315 GTGGATTTTAAACAGATATTTTATTTTATTTATTTGAGACAAATGTTGA 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65662 GTGGATTTTGA---AGATATCTATTTTATTTATTTGAGACAAATGTTGA 65613

```

```

RESULT 13
AC119115 138792 bp DNA linear HTG 18-JUL-2002
LOCUS AC119115
DEFINITION Rattus norvegicus clone CH230-320N23.*** SEQUENCING IN PROGRESS
***, 32 unordered pieces.

```

## ACCESSION

AC119115  
AC119115.2 GI:21746718  
HTG: HTGS\_PHASE1.  
Norway rat.  
SOURCE  
Rattus norvegicus  
Eukaryota; Metazoa;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

1 (bases 1 to 138792)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshman, F.R., Allen, C., Albrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Blum, K., Blankenburg, K., Bonin, D., Bouck, J., Boyle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, N., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, S., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, R., Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Medora, M., Mei, G., Melker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J.T., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S., Ogih, M., Okunoda, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojudo, A., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tameris, A., Tameris, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlecczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

## TITLE

Unpublished  
2 (bases 1 to 138792)

## REFERENCE

Worley, K.C.  
Direct Submission  
Submitted (25-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 138792)

## REFERENCE

Worley, K.C.  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20303440.

## JOURNAL

Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc.helpebcm.tmc.edu  
Project Information  
Center project name: G0XG  
Center clone name: CH230-320N23  
Summary Statistics

## COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc.helpebcm.tmc.edu  
Project Information  
Center project name: G0XG  
Center clone name: CH230-320N23  
Summary Statistics





	AC130192/c			
LOCUS	AC130192	177555 bp	DNA	linear HTG 08-AUG-2002
DEFINITION	Sus scrofa clone RP44-436R21, WORKING DRAFT SEQUENCE, 10 unordered pieces.			
ACCESSION	AC130192			
VERSION	AC130192.1 GI:22138443			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE	pig.			
ORGANISM	Sus scrofa			
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia: Eutheria; Cetartiodactyla; Suidae; Sus.			
AUTHORS	Mammals: 1 (bases 1 to 177555) Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B., Maguires,E.H., Mastello,C., Maskeri,B., Mastilan,S.D., McLoiskey,J.C., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Scheller,M.G., Stantiford,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A., Weheby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D. NSC Comparative Sequencing Initiative Unpublished 2 (bases 1 to 177555)			
JOURNAL TITLE	Direct Submission			
JOURNAL REFERENCE	Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717 Goveoment Circle, Gaithersburg, MD 20877, USA			
TITLE COMMENT	----- Genome Center			

```

----- Summary Statistics -----
Center clone name: 436K21
Sequencing vector: plasmid; N/A; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164120 bases at least Q40
Consensus quality: 168083 bases at least Q30
Consensus quality: 170767 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 176655; sum-of contigs
Quality coverage: 4.78x in Q20 bases; agarose-fp
Quality coverage: 4.57x in Q20 bases; sum-of contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 3461 3560: contig of 3460 bp in length
* 3561 6669: contig of 3309 bp in length
* 6870 6669: gap of unknown length
* 6970 13465: contig of 6496 bp in length
* 13466 13665: gap of unknown length
* 13666 21301: contig of 7736 bp in length
* 21302 21301: gap of unknown length
* 21402 33719: contig of 12318 bp in length
* 33720 33819: gap of unknown length
* 33820 45054: contig of 11235 bp in length
* 45055 45154: gap of unknown length
* 45155 58724: contig of 13570 bp in length
* 58725 58824: gap of unknown length
* 58825 76670: contig of 17846 bp in length
* 76671 76670: gap of unknown length

```

FEATURES	SOURCE
* 76771	115165: contig of 38395 bp in length
* 115166	115265: gap of unknown length
* 115266	177555: contig of 62290 bp in length
	Location/Qualifiers
	1. 177555

```

misc_feature      1.3460
                   /note="assembly-fragment
misc_feature      3551.6869
                   /note="assembly-fragment
misc_feature      6970.13465
                   /note="assembly-fragment"
misc_feature      13566.21301
                   /note="assembly-fragment"
misc_feature      21402.33719
                   /note="assembly-fragment"
misc_feature      33820.45054
                   /note="assembly-fragment"
misc_feature      43515.58724
                   /note="assembly-fragment
clone_end:17
vector_side:left"
misc_feature      58825.76670
                   /note="assembly-fragment"
misc_feature      76771.115165
                   /note="assembly-fragment"
misc_feature      115266.177555

```

BASE COUNT	ORIGIN
41026 a	46997 c 46681 g 41924 t 927 others

Query Match	17.9%;	Score 246.4;	DB 2;	Length 177555;
Best Local Similarity	85.6%;	Pred. No. 3.2e-35;		
Matches 274;	Conservative	0;	Mismatches 46;	Indels 0; Gaps 0;

QY	591	CCCTACTGTCAGGTGCACCTTATATGAGGGGAAGGCGTCTACTCTGAAGCTGACCTGCT	650
Db	114241	CCCCCCCCCAGGTGCACCTTATATGAGGGGAAGGCGTCTACTCTGAAGCTGACCTGCT	114182
QY	651	GGTGGATGCTGTGCTGGCCCTTGCCGCTGCCTCGAGAGAAATTCACAGCCACTGCGGCCAGTTC	710
Db	114181	GGTGGAGCAGACACGCTGGCCCTTGCCGCTGCCTCGAGAGAAATTCACAGCCAGCGCGCAGATTC	114122
QY	711	CCCTGGGGCCCCAGCTCCGCGCTCTGCCAGAGGTGTCTGGGCTCTTGGCCCTGGCGGCCAGAGTTC	770
Db	114121	CCCTGGGGCCCCAGCTCCGCGCTCTGCCAGAGGTGTCTGGGCTCTTGGCCCTGGCGGCCAGAGTTC	114062
QY	771	CTCCCTTGCGGATTCGCGACCCCTCCCTGGGGCCATCTCAGCAAGGCTGCCCTTCTCTCACTTA	830
Db	114061	CTCCCTTGCGGATTCGCGACCCCTCCCTGGGGCCACATCAAGGCTGCCCTTCTCTCACTTA	114002
QY	831	CTTGGGACTCTTCAGAGTCTACGTAAGGGGGCCCTGGTCTCCCAAGTCGTCGCCAGGCTCG	890
Db	114001	CTTGGGACTCTTCAGAGTCTACGTAAGGGGGCCCGGGGCTCCCAAGTCGTCGCCAGGCTCCACCCG	113942
QY	891	CGGCTCCCTCGACAGCTCT	910
Db	113941	GCGTGCCTCTCCCGAGACT	113922

RESULT 15  
AC126925/c LOCUS  
DEFINITION  
AC126925 Canis familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 15  
uncloned pieces.  
AC126925  
VERSION AC126925.1 GI:21724102  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.



Db	42211	CTCAAGCTCCCCCAACCCCCCTGCCTGACCTGCCCTCTGACAGGCTGCTGGACTTG	42152
QY	961	CTGCCCCCTCCCTAGAGGCTGCCGTGGCCCTGTTACAGTGTGTTTCCATCCCATTAAT	1020
Db	42151	TTCAACCTGTTTCGGGACACATACATATATCTCTCCCTCTGACACCCCATCT	42096
QY	1021	ACAGTATCCCACTCTTATCTTACACATGCCCCACCGCCACTCTCCACTCATTAGCTC	1080
Db	42095	-----CCACGTCACCTGGGCCCTGAACCCCTGCTTTTGAAGCCCCCAATTATCTC	42045
QY	1081	CCCAATCCCTGACCCCTTTGAGGCCCCCAGTGTCTGCACCTCCCCCTGGCCACAGACCC	1140
Db	42044	CTGACTTCTTAATCTTGTGGCAACCCCGCCACCCCTCCCTCCCCCAACACACAGC----	41989
QY	1141	CAGGCATTGTGTTACTGTACTCTGTGTGGGCAAGATGGGTTCAGAA---GACCCCACTT	1197
Db	41988	-----CAAAAGCACATGTGTCTATCTGTGGTGAATGGGTCTCCCAACCTCCCCCACTT	41934
QY	1198	CAGGACATAAGAGGGGCTGGAGCTGGCCGGCAGAGAACCAAGAGACTGGGCTTAAGCCAG	1257
Db	41933	CAGGACATAATCCAGGGCTAGAGGCCAGGGGCGCACTGGGACAG-----GAAGCCGG	41884
QY	1258	GAGTCCCAAAATGTAGGGGGAGAGAAACAAGCAAGCTCCCTCCCTTGAAGAAATCCCTGTG	1317
Db	41883	GAGTCCCAATACGCGCAGGGCTGAG-AACAAGGTGAGCTCC-----TTCCCTGTG	41836
QY	1318	GATTTTAAAAACAGATATTTATTTTATTATTATTTGTGACAAAAATGTGTGATAATGG	1373
Db	41835	GATTTTAAAAACAGATATTTATTTTATTATTATTTGTGACAAAGATCTCAATTAAGTGG	41780

Search completed: May 8, 2003, 10:02:32  
Job time : 4943.14 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 11:47:25 ; Search time 1852 Seconds  
(without alignments)  
12006.697 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373

Sequence: 1 atgcattgttagacttga.....gacaaatgttgaataatg 1373

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estinu:\*  
4: em\_estov:\*  
5: em\_estov:\*  
6: em\_estov:\*  
7: em\_estov:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estfun:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	45.8	777	13	BI819200
2	613	44.6	1071	14	BM921213
3	605	44.1	963	14	BM671259
4	518	37.7	828	13	BI596681
5	506	36.9	728	13	BI870393
6	506	36.9	731	13	BI871711

Result No.	Score	Query Match	Length	DB ID	Description
1	629	45.8	777	13	BI819200
2	613	44.6	1071	14	BM921213
3	605	44.1	963	14	BM671259
4	518	37.7	828	13	BI596681
5	506	36.9	728	13	BI870393
6	506	36.9	731	13	BI871711

# ALIGNMENTS

RESULT 1  
BI819200  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BI819200 777 bp mRNA linear EST 04-OCT-2001  
603034614P1 NIH\_MGC\_115 Homo sapiens cdna clone IMAGE:5175698 5',  
mRNA sequence.  
BI819200  
BI819200.1 GI:15930750  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 777)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-ri@mail.nih.gov](mailto:cgabs-ri@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.llnl.gov>  
Plate: LLM11437 row: 1 column: 03  
High quality sequence stop: 759.  
Location/Qualifiers  
1..777

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/clone_id="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      140 a      268 c      215 g      154 t
ORIGIN

```

```

Query Match      45.8%; Score 629; DB 13; Length 777;
Best Local Similarity 99.7%; Pred. No. 4.8e-313;
Matches 729; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 288 GGCAGAGAGACACAGACCCGCTGGAATCCCAAGACAGAAAGCCAGATCC 347
    |||
Db 13 GGCAGAGAGACACAGACCCGCTGGAATCCCAAGACAGAAAGCCAGATCC 72

QY 348 TGGCGCTTCCGACACCACTAGTGGGCTCGCAAGAGTGACCTAAAGCCGGAATC 407
    |||
Db 73 TGGCGCTTCCGACACCACTAGTGGGCTCGCAAGAGTGACCTAAAGCCGGAATC 132

QY 408 ACGGCTCGAAGAGCATGCGACGCCATTATGAATCATCCAGACCTGGACAGGAGG 467
    |||
Db 133 ACGGCTCGAAGAGCATGCGACGCCATTATGAATCATCCAGACCTGGACAGGAGG 192

QY 468 AGCGAGGACAGGTGAGAGGAGAGTGGCTGGAGAGAGCAGAAATCAACAGCTC 527
    |||
Db 193 AGCGAGGACAGGTGAGAGGAGAGTGGCTGGAGAGAGCAGAAATCAACAGCTC 252

QY 528 CAGGCTGTGGCTACAAACCCGAGATGGGAGTTTATGACACCGGCTGGCTCTA 587
    |||
Db 253 CAGGCTGTGGCTACAAACCCGAGATGGGAGTTTATGACACCGGCTGGCTCTA 312

QY 588 CTACCTGTACTGTAGAGTGACACTTGTATGAGGAGAGGCTGTACTACCTGAAGCTGGAAT 647
    |||
Db 313 CTACCTGTACTGTAGAGTGACACTTGTATGAGGAGAGGCTGTACTACCTGAAGCTGGAAT 372

QY 648 GCTGTGTAGTGTGTGCTGCGCTGCGCTGAGAGAAATTCAGGCACTGCGGCGAG 707
    |||
Db 373 GCTGTGTAGTGTGTGCTGCGCTGCGCTGAGAGAAATTCAGGCACTGCGGCGAG 432

QY 708 TTCCCTGGGGCCGAGCCGCTGCTGCAAGTGTCTGGCTGTGGCCCTGGCGCCAGG 767
    |||
Db 433 TTCCCTGGGGCCGAGCCGCTGCTGCAAGTGTCTGGCTGTGGCCCTGGCGCCAGG 492

QY 768 GTCTCCCTGCGGATCCGACACCTCCCTGGGCCCATTCTAAAGCTGCGCCCTTCTCTAC 827
    |||
Db 493 GTCTCCCTGCGGATCCGACACCTCCCTGGGCCCATTCTAAAGCTGCGCCCTTCTCTAC 552

QY 828 CTACTTGGGACTCTTCAAGTTCACTAGAGGGCCCTGCTGCCACAGTGTCCAGGC 887
    |||
Db 553 CTACTTGGGACTCTTCAAGTTCACTAGAGGGCCCTGCTGCCACAGTGTCCAGGC 612

QY 888 TGGCGGCTCCGACAGCTCTGAGGACCCGCTGCCCTCTGCCCCCAGCCCTAGCCGC 947
    |||
Db 613 TGGCGGCTCCGACAGCTCTGAGGACCCGCTGCCCTCTGCCCCCAGCCCTAGCCGC 672

QY 948 TCTTGTCTCAGACCTCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTTTTC 1007
    |||
Db 673 TCTTGTCTCAGACCTCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTTTTC 732

QY 1008 ATCCGACATTA 1018
    |||
Db 733 ATCCGACATTA 743

```

```

RESULT 2
BM921213
LOCUS
DEFINITION
BM921213 1071 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:51752561
5', mRNA sequence.
ACCESSION
BM921213
VERSION
BM921213.1 GI:19371592
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1071)
NIH-MGC http://mgc.ncl.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12786 row: P column: 02
High quality sequence stop: 656.
location/Qualifiers
1. 1071
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:51752561"
/clone_id="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      226 a      346 c      279 g      220 t
ORIGIN

```

```

Query Match      44.6%; Score 613; DB 14; Length 1071;
Best Local Similarity 99.7%; Pred. No. 8.1e-305;
Matches 713; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 412 GCTGAGAGAGGATGAGACGCCATTATGAAGTTATCCAGACCTGGACAGGAGCG 471
    |||
Db 12 GCTGAGAGAGGATGAGACGCCATTATGAAGTTATCCAGACCTGGACAGGAGCG 71

QY 472 CAGGACAGTGTGAGAGGAGAGTGTGCTGGAGAGAGCCAGAAATCAACAGCTCCAGC 531
    |||
Db 72 CAGGACAGTGTGAGAGGAGAGTGTGCTGGAGAGAGCCAGAAATCAACAGCTCCAGC 131

QY 532 CTTCTGCGTACAAACCCGAGATCGGGAGTTATAGTCAACCGGCTGGGCTCTACTAC 591
    |||
Db 132 CTTCTGCGTACAAACCCGAGATCGGGAGTTATAGTCAACCGGCTGGGCTCTACTAC 191

QY 592 CTTTACTGTAGAGTGCATTGTATGAGGAGAGGAGCTCTTACTGAAAGCTGAGTGTCTG 651
    |||
Db 192 CTTTACTGTAGAGTGCATTGTATGAGGAGAGGAGCTCTTACTGAAAGCTGAGTGTCTG 291

QY 652 GTGAGTGTGTGTGCGCTGCGCTGCGCTGAGGAATTCACGCCACTGGCGGAGTTCC 711
    |||
Db 252 GTGAGTGTGTGTGCGCTGCGCTGCGCTGAGGAATTCACGCCACTGGCGGAGTTCC 311

```



```

OY 712 CTCGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTGGCGCCAGGGTCC 771
DB 312 CTCGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTGGCGCCAGGGTCC 371
OY 772 TCCCTGGGATCGGACACCTCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCTCTACCTAC 831
DB 372 TCCCTGGGATCGGACACCTCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCTCTACCTAC 431
OY 832 TTCGGACTTTCAGGTTACTAGAGGGGCGCTGGTCTCCCGACAGCTGCCAGGGTCC 891
DB 432 TTCGGACTTTCAGGTTACTAGAGGGGCGCTGGTCTCCCGACAGCTGCCAGGGTCC 491
OY 892 GGCCTCCCTCGACAGCTCTCTGCGGACCCGCTCCCTCTGCGCCACCTCTAGCGCTCTT 951
DB 492 GGCCTCCCTCGACAGCTCTCTGCGGACCCGCTCCCTCTGCGCCACCTCTAGCGCTCTT 551
OY 952 TGTCTCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTCTACGCTTTTTCATCC 1011
DB 552 TGTCTCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTCTACGCTTTTTCATCC 611
OY 1012 CACATTAATACAGTATTTCCACTCTATCTTACAACTCCCGACGCGCACTCTCCAGCT 1071
DB 612 CACATTAATACAGTATTTCCACTCTATCTTACAACTCCCGACGCGCACTCTCCAGCT 671
OY 1072 CACTAGCTCCCGCAATCCCTGACCCCTTGGAGGCCCGCCAGTATCTGACTCCGCC 1126
DB 672 CACTAGCTCCCGCAATCCCTGACCCCTTGGAGGCCCGCCAGTATCTGACTCCGCC 726

```

```

RESULT 3
BO671259 963 bp mRNA linear EST 15-JUL-2002
LOCUS BO671259
DEFINITION AGENCORT.8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
5', mRNA sequence.
VERSION BO671259
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 963)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2456 row: 1 column: 13
High quality sequence stop: 565.
Location/Qualifiers
1..963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6274716"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: salivary gland; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

```

BASE COUNT

194 a 326 c 260 g 183 t

## .ORIGIN

Query Match 44.1%; Score 605; DB 14; Length 963;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-300;  
 Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 408 ACGGGCTCAAGAGCATGACAGCCATTATGAGTTCAATCCACGACCTGGACAGAGAGG 467
DB 39 ACGGGCTCAAGAGCATGACAGCCATTATGAGTTCAATCCACGACCTGGACAGAGAGG 98
OY 468 ACGGACGAGAGTGTGAGGAGGAGAGTGAAGTGGCTGGAGGAGGAGGAGGAGGAGG 527
DB 99 ACGGACGAGAGTGTGAGGAGGAGAGTGAAGTGGCTGGAGGAGGAGGAGGAGGAGG 158
OY 528 CAGCCCTCTGCGCTACAAACCGGCAATCGGGAGTTATAGTCACCCGGGCTGGCTCTA 587
DB 159 CAGCCCTCTGCGCTACAAACCGGCAATCGGGAGTTATAGTCACCCGGGCTGGCTCTA 218
OY 588 CTACCTGTACTGTCTAGTGTCTAGTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 647
DB 219 CTACCTGTACTGTCTAGTGTCTAGTGTATGAGGAGGAGGAGGAGGAGGAGGAGG 278
OY 648 GCTGTGTGATGTTGTGCTGAGGCTGCGCTGCTGAGGAGGAGGAGGAGGAGGAGG 707
DB 279 GCTGTGTGATGTTGTGCTGAGGCTGCGCTGCTGAGGAGGAGGAGGAGGAGGAGG 338
OY 708 TTTCCCTCGGGCCCCAGCTCCGCTCTGCGGAGTGTCTGGGCTGTGGCTTGGCCCTG 767
DB 339 TTTCCCTCGGGCCCCAGCTCCGCTCTGCGGAGTGTCTGGGCTGTGGCCCTGCGG 398
OY 768 GTCTCTCCCTGGGAGATCCGACACCTCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 827
DB 399 GTCTCTCCCTGGGAGATCCGACACCTCCCTGGGAGGAGGAGGAGGAGGAGGAGG 458
OY 828 CTACTTCTGAGCTCTTCTCAGTCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 887
DB 459 CTACTTCTGAGCTCTTCTCAGTCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 518
OY 888 TGGCGGCTCCCTCGACACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 947
DB 519 TGGCGGCTCCCTCGACACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 578
OY 948 TCTTCTCTCAGACCTGCGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1007
DB 579 TCTTCTCTCAGACCTGCGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 638
OY 1008 ATCCC 1012
DB 639 ATCCC 643

```

```

RESULT 4
LOCUS B1596681 828 bp mRNA linear EST 07-SEP-2001
DEFINITION 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
B1596681 mRNA sequence.
VERSION B1596681
KEYWORDS EST.
SOURCE B1596681.1 GI:15489620
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 828)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

```

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>

Plate: L1AM1722 row: k column: 13  
 High quality sequence stop: 776.  
 Location/Qualifiers

## FEATURES

source

1. 828  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5285892"  
 /clone\_1lb="NIH\_MGC\_96"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTT-3',  
 size-selected for average insert size 2.3 kb and  
 normalized to 10<sup>5</sup>. This is a primary library enriched  
 for full-length clones and constructed using the  
 CAP-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 155 a 278 c 223 g 172 t  
 ORIGIN

Query Match 37.7%; Score 518; DB 13; Length 828;  
 Best Local Similarity 99.7%; Pred. No. 8.2e-256;

Matches 618; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

312 GGAAGTGAATCCCGACAGAGAAAGCAGATCCCGCTTCTGAAACGACTAGT 371  
 |||||  
 43 GGAATGAATCCCGACAGAGAAAGCAGATCCCGCTTCTGAAACGACTAGT 102  
 |||||  
 372 TCGGCTCGAGAGAGTCCACTTAAAGCCGGAACACAGGCTCGAAGAGGATCGAGC 431  
 |||||  
 103 TCGGCTCGAGAGTCCACTTAAAGCCGGAACACAGGCTCGAAGAGGATCGAGC 162  
 |||||  
 432 CCATTATGAAGTTATCCAGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 491  
 |||||  
 163 CCATTATGAAGTTATCCAGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 222  
 |||||  
 492 AGTAGTGGTGGAG 551  
 |||||  
 223 ATTGAGTGGTGGAG 282  
 |||||  
 552 GATCGGGAGATTATAGTCAACCGGGCTGAGCTTACTACTGTAAGTGTGACTT 611  
 |||||  
 283 GATCGGGAGATTATAGTCAACCGGGCTGAGCTTACTACTGTAAGTGTGACTT 342  
 |||||  
 612 TGATGAGGAGAGAGCTGTCTACTGAGAGTGTGAGTGTGAGTGTGAGTGTG 671  
 |||||  
 343 TGATGAGGAGAGAGCTGTCTACTGAGAGTGTGAGTGTGAGTGTGAGTGTG 402  
 |||||  
 672 GCGCTGCTGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 731  
 |||||  
 403 GCGCTGCTGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 462  
 |||||  
 732 CTGCGAGGTGTGAGGCTGTGAGGCTGCGGCAAGGCTTCCCTGGGATTCGCA 791  
 |||||  
 463 CTGCGAGGTGTGAGGCTGTGAGGCTGCGGCAAGGCTTCCCTGGGATTCGCA 522  
 |||||  
 792 CCGCTGGGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 851  
 |||||  
 523 CCGCTGGGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 582  
 |||||  
 852 CTGAGGAGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 911  
 |||||  
 583 CTGAGGAGAGAGATTTCAAGCACTGGGGAGTTCCTGGGGCCCAAGTCCGCT 642  
 |||||  
 912 TGGGCAACCGGCTCCCTCTG 931  
 |||||  
 643 TGGGCAACCGGCTCCCTCTG 662  
 |||||

RESULT 5  
 BI870393 728 bp mRNA linear EST 11-OCT-2001  
 LOCUS 60339564.F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5405459.5,  
 DEFINITION mRNA sequence.

ACCESSION BI870393.1 GI:16044066  
 VERSION BI870393  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 728)

NIH-MGC <http://imgc.ncl.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cga@rs-remail.nih.gov](mailto:cga@rs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1AM12034 row: a column: 12

High quality sequence stop: 728.

## FEATURES

source

Location/Qualifiers

1. 728

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5405459"

/clone\_1lb="NIH\_MGC\_90"

/tissue\_type="adenoecarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; Oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 125 a 240 c 227 g 136 t

## ORIGIN

Query Match 36.9%; Score 506; DB 13; Length 728;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-249;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

262 CAGGAGCTGCTCCCGACGAGAGTGTGCGACAGAGAGAGAGAGAGAGAGAG 321  
 |||||  
 74 CAGGAGCTGCTCCCGACGAGAGTGTGCGACAGAGAGAGAGAGAGAGAGAG 133  
 |||||  
 322 CCCGACAG 381  
 |||||  
 134 CCCGACAG 193  
 |||||  
 382 AGAAGTCACTTAAAGCCGGAAGACAGGAGTGTGAGAGAGAGAGAGAGAGAG 441  
 |||||  
 194 AGAAGTCACTTAAAGCCGGAAGACAGGAGTGTGAGAGAGAGAGAGAGAGAG 253  
 |||||  
 442 GTTCATCCAG 501  
 |||||  
 254 GTTCATCCAG 313  
 |||||  
 502 TGGGAG 561  
 |||||  
 314 TGGGAG 373  
 |||||  
 562 TTATATGATCACCGGGCTGAGCTTACTACTTACTAGTGTGAGAGAGAGAGAG 621  
 |||||  
 374 TTATATGATCACCGGGCTGAGCTTACTACTTACTAGTGTGAGAGAGAGAGAG 433  
 |||||

```

OY 622 AAGCGTGTACCTGAAGCTGAGCTGCTGTGTGATGTGTGTGCTGGCCCTGGCCTGCTG 681
    |||||||
DB 434 AAGCGTGTACCTGAAGCTGAGCTGCTGTGTGATGTGTGTGCTGGCCCTGGCCTGCTG 493
OY 682 GAGGAATTCACAGCCACTGCGGCGACATTCCTCGGGCCCCAGCTCCGCTTGCAGGTG 741
    |||||||
DB 494 GAGGAATTCACAGCCACTGCGGCGACATTCCTCGGGCCCCAGCTCCGCTTGCAGGTG 553
OY 742 TCTGGGCTGTGGCCCTGGCGCCAGG 767
    |||||||
DB 554 TCTGGGCTGTGGCCCTGGCGCCAGG 579

RESULT 6
BI871711 731 bp mRNA linear EST 11-OCT-2001
LOCUS DEFINITION 603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
    mRNA sequence.
ACCESSION BI871711
VERSION BI871711.1 GI:16045386
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
    1 (bases 1 to 731)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
    Email: cga@bbs-riemail.nih.gov
COMMENT Tissue Procurement: ATCC
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/ILNI at:
    http://image.llnl.gov
    Plate: L1M12034 row: b column: 07
    High quality sequence stop: 728.
FEATURES
    source
        1..731
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5405478"
            /clone_lib="NIH_MGC_90"
            /tissue_type="adenocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Liver; Vector: pCMV-SPORT6; Site.1: NotI;
            Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 1.7 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH-MGC Library."
BASE COUNT 125 a 242 c 228 g 136 t
ORIGIN
Query Match 36.9%; Score 506; DB 13; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.3e-249;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 262 CAGGAGCTGCGCCAGAGAGCTGTGTGAGAGAGGAGGACGACCCCGCGAAGCAAT 321
    |||||||
DB 74 CAGGAGCTGCGCCAGAGAGAGCTGTGTGAGAGAGGAGGACGACCCCGCGAAGCAAT 133
OY 322 CCCAGACAGAAAGAACCCAGATCTGCGCTTCTCCTGAACGACTAGTGGCTCGC 381
    |||||||
DB 134 CCCAGACAGAAAGAACCCAGATCTGCGCTTCTCCTGAACGACTAGTGGCTCGC 193
OY 382 AATAATGCACTTAAGGCGCGGAAACAGCGGCTCGAAGCGATGCGACCACTTATGA 441
    |||||||
DB 194 AATAATGCACTTAAGGCGCGGAAACAGCGGCTCGAAGCGATGCGACCACTTATGA 253
OY 442 GTTCATCCAGACTGAGACGAGGAGGCGAGGAGGTGTGAGGAGGACATGAGTGGC 501

```

```

DB 254 GTTCATCCAGACTGAGACAGGAGGAGGAGGAGGTGTGAGGAGGACATGAGTGGC 313
    |||||||
OY 502 TGGAGAGAGCCAGAAATCAACAGCTTCAGCCCTTGCCTACAAACCCAGATCGGAG 561
    |||||||
DB 314 TGGAGAGAGCCAGAAATCAACAGCTTCAGCCCTTGCCTACAAACCCAGATCGGAG 373
OY 562 TTTATATGACCCCGGCGGCGCTCTACCTGCTACTGTACGTGACATTTATGAGGG 621
    |||||||
DB 374 TTTATATGACCCCGGCGGCGCTCTACCTGCTACTGTACGTGACATTTATGAGGG 433
OY 622 AAGCGTGTACCTGAAGCTGAGCTGCTGTGTGATGTGTGTGCTGGCCCTGGCCTGCTG 681
    |||||||
DB 434 AAGCGTGTACCTGAAGCTGAGCTGCTGTGTGATGTGTGTGCTGGCCCTGGCCTGCTG 493
OY 682 GAGGAATTCACAGCCACTGCGGCGACATTCCTCGGGCCCCAGCTCCGCTTGCAGGTG 741
    |||||||
DB 494 GAGGAATTCACAGCCACTGCGGCGACATTCCTCGGGCCCCAGCTCCGCTTGCAGGTG 553
OY 742 TCTGGGCTGTGGCCCTGGCGCCAGG 767
    |||||||
DB 554 TCTGGGCTGTGGCCCTGGCGCCAGG 579

RESULT 7
BI966060 609 bp mRNA linear EST 12-MAR-2002
LOCUS DEFINITION 1e72904.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
    cDNA clone IMAGE:5672623 3' similar to TR:054907 054907 TNF-RELATED
    WEAK INDUCER OF APOPTOSIS 1, mRNA sequence.
ACCESSION BI966060
VERSION BI966060.1 GI:16340465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS Melton, D., Brown, J., Kenty, G., Penmutt, A., Lee, C., Kaestner, K.,
    Lemishka, T., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
    Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A.,
    Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
    , M., Gibbons, M., McCann, R., Cole, R., Tsagarisvill, R., Williams, T.,
    Jackson, Y., and Bowers, Y.
    Endocrine Pancreas Consortium
    Unpublished (2000)
    Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
    Endocrine Pancreas Consortium
    Harvard University, Howard Hughes Medical Institute
    Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
    MA 02138
    Tel: 617-495-1812
    Fax: 617-495-8557
    Email: dmelton@biohp.harvard.edu
    Library was constructed by Dr. Douglas Melton DNA sequencing by:
    Washington University Genome Sequencing Center For information on
    obtaining a clone please contact: Juliana Brown
    (brownjfas.harvard.edu) This sequence now available from the IMAGE
    consortium, for clone orders contact: info@image.llnl.gov
    High quality sequence stop: 412.
FEATURES
    source
        1..609
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5672623"
            /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
            /sex="Both"
            /tissue_type="Islets of Langerhans"
            /dev_stage="Adult"
            /lab_host="DH10B"
            /note="Organ: Pancreas; Vector: pSPORT1; Site.1: Not 1;
            Site.2: Sal 1; Starting library constructed using
            Superscript plasmid library kit (Life Technologies). cDNA

```

made by oligo-dt priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and 1.08 kb DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoI of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others  
ORIGIN

Query Match 35.8%; Score 491; DB 13; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.9e-242;  
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 CAGCTGCGGCTCCCTCGACAGCTCTGTGGCAGCCGGTCCCTGCGCCACCTCA 942  
|||||  
DB 513 CAGCTGCGGCTCCCTCGACAGCTCTGTGGCAGCCGGTCCCTGCGCCACCTCA 454  
943 GCGCTCTTTGGTCCAGACCTGCCCCCTCTAGAGGCTGCGCTGTTACAGTGT 1002  
DB 453 GCGCTCTTTGGTCCAGACCTGCCCCCTCTAGAGGCTGCGCTGTTACAGTGT 394  
OY 1003 TTTCCATCCCATATAAATACAGTATCCACCTTATCTTACACTCCCCACGCCAC 1062  
|||||  
DB 393 TTTCCATCCCATATAAATACAGTATCCACCTTATCTTACACTCCCCACGCCAC 334  
OY 1063 TCTCCACCTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGATGTGACATC 1122  
DB 333 TCTCCACCTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGATGTGACATC 274  
OY 1123 CCCCTGGCCAGACCCCCCAGGCAATTTGTCTACTGACTCTGTGGCAGAGATGGTTC 1182  
DB 273 CCCCTGGCCAGACCCCCCAGGCAATTTGTCTACTGACTCTGTGGCAGAGATGGTTC 214  
OY 1183 CAGAGACCCGACCTTACAGCACTAGAGGGCTGACCTGCGGCGGAGAGCAAGAGA 1242  
DB 213 CAGAGACCCGACCTTACAGCACTAGAGGGCTGACCTGCGGCGGAGAGCAAGAGA 154  
OY 1243 CTGGGCTTAGGCGCAGAGATTCCTCAATGTGAGGGCGAGAAACAAGACCTCTCCT 1302  
DB 153 CTGGGCTTAGGCGCAGAGATTCCTCAATGTGAGGGCGAGAAACAAGACCTCTCCT 94  
OY 1303 TGAGATTCCTCTGTGGATTTTAAACAGATATTATTTATTTATTTATTTGACAAATG 1362  
DB 93 TGAGATTCCTCTGTGGATTTTAAACAGATATTATTTATTTATTTATTTGACAAATG 34  
OY 1363 TTGATTAATGG 1373  
DB 33 TTGATTAATGG 23

RESULT 8 948 bp mRNA linear EST 16-JUL-2002  
B0707185  
LOCUS AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6278608  
DEFINITION 5', mRNA sequence.  
ACCESSION B0707185  
VERSION B0707185.1 GI:21846084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 948)  
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@pds-remail.nih.gov

Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.nih.gov  
Plate: L10M2466 row: n column: 17  
High quality sequence start: 24  
High quality sequence stop: 550.

FEATURES  
SOURCE

location/qualifiers

1..948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6278608"  
/clone\_11b="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Query Match 31.8%; Score 437; DB 14; Length 948;  
Best Local Similarity 99.8%; Pred. No. 5.1e-214;  
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 475 GCAGGTGTGACGGGACAGTGTGCTGGAGAGGACAGAAATCAACAGCTCCAGCCCT 534  
DB 167 GCAGGTGTGACGGGACAGTGTGCTGGAGAGGACAGAAATCAACAGCTCCAGCCCT 226  
OY 535 CTGGCTCAACCCGCAATGTGGGAGATTATATGACACCCGGGCTGGCTTACTACTG 594  
DB 227 CTGGCTCAACCCGCAATGTGGGAGATTATATGACACCCGGGCTGGCTTACTACTG 286  
OY 595 TACTGTCAAGTCACTTGTATGATGAGGAGAGCTGTCTACTGAGCTGAGTCTGTG 654  
DB 287 TACTGTCAAGTCACTTGTATGATGAGGAGAGCTGTCTACTGAGCTGAGTCTGTG 346  
OY 655 GATGTGTGTGGCCCTGCTGCTGTGAGGAAATTTCTAGCCACTGCGGCGCAAGTTCCCTC 714  
DB 347 GATGTGTGTGGCCCTGCTGCTGTGAGGAAATTTCTAGCCACTGCGGCGCAAGTTCCCTC 406  
OY 715 GGGCCCGACGTCGCTGTGCAAGTGTCTGCTGTGCGCTGTGCGCCGCGGAGGCTCTCC 774  
DB 407 GGGCCCGACGTCGCTGTGCAAGTGTCTGCTGTGCGCTGTGCGCCGCGGAGGCTCTCC 466  
OY 775 CTGGGATCCGACCTCCCTGCGCCATCTCAAGCTGCGCCCTTCTCACTCACTTTC 834  
DB 467 CTGGGATCCGACCTCCCTGCGCCATCTCAAGCTGCGCCCTTCTCACTCACTTTC 526  
OY 835 GCAGCTTTCAGGTTCACTAGAGGCGCTGTCTCCCAAGTGTGCTCCAGGCTGCGGC 894  
DB 527 GCAGCTTTCAGGTTCACTAGAGGCGCTGTCTCCCAAGTGTGCTCCAGGCTGCGGC 586  
OY 895 TCCCTGACAGCTCTGTGGGACCGGCTCCCTGCTCCCAAGCTGAGCGGCTTTCG 954  
DB 587 TCCCTGACAGCTCTGTGGGACCGGCTCCCTGCTCCCAAGCTGAGCGGCTTTCG 646  
OY 955 TCCAGACC 962  
DB 647 TCCAGACC 654

RESULT 9 940 bp mRNA linear EST 16-AUG-2002  
B0884231  
LOCUS B0884231  
DEFINITION AGENCOURT\_8682031 Lupscl\_sciatic nerve Homo sapiens cDNA clone

IMAGE:6197488 5' mRNA sequence.  
 ACCESSION B0884231  
 VERSION B0884231.1 GI:22276239  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 940)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
 Plate: L1AM13607 row: j column: 17  
 High quality sequence stop: 453.

FEATURES  
 source  
 1. 940  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="6197488"  
 /clone\_lib="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-Sport6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGGCG-3' and 5'-GACTAGTCTAGATCGAGCGGCGGCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN  
 Query Match 31.8%; Score 436; DB 14; Length 940;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-213;  
 Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 301 CAGGACCGGTGGAGCTAAATCCAGACAGAAAGCCAGATCTGGCCTTCTG 360  
 DB 1 CAGGACCGGTGGAGCTAAATCCAGACAGAAAGCCAGATCTGGCCTTCTG 60  
 OY 361 AACGCTAGTTCGGCTTCGAGAAATGACCTAAAGCCGGAACACAGGCTCGAAGA 420  
 DB 61 AACGCTAGTTCGGCTTCGAGAAATGACCTAAAGCCGGAACACAGGCTCGAAGA 120  
 OY 421 GCGATCGACGCCATTAATGAACTTCACAGACCTGAGACGAGACGAGCAGAGT 480  
 DB 121 GCGATCGACGCCATTAATGAACTTCACAGACCTGAGACGAGACGAGCAGAGT 180  
 OY 481 GTGAGCGGAGAGTGTGGCTGGAGAGCAATCAAGTCCAGCCCTCTGGC 540  
 DB 181 GTGAGCGGAGAGTGTGGCTGGAGAGCAATCAAGTCCAGCCCTCTGGC 240  
 OY 541 TACAACCCGACATCGGGAGTTTATAGTCAACCCGGGCTGGCTACTACTACTGT 600  
 DB 241 TACAACCCGACATCGGGAGTTTATAGTCAACCCGGGCTGGCTACTACTACTGT 300  
 OY 601 CAGGTGCACTTGTATGAGAGGAGCTGTCTACTGAAAGCTGACTTGTGATGATGT 660  
 DB 301 CAGGTGCACTTGTATGAGAGGAGCTGTCTACTGAAAGCTGACTTGTGATGATGT 360

OY 661 GTGCTGCCCTCGCTGCTGAGGAATTCACGACTGCGGCGAGTTCCTCGGCCCC 720  
 DB 361 GTGCTGCCCTCGCTGCTGAGGAATTCACGACTGCGGCGAGTTCCTCGGCCCC 420  
 OY 721 CAGTCCGCTCTGCGC 736  
 DB 421 CAGTCCGCTCTGCGC 436

RESULT 10  
 LOCUS B1766766 834 bp mRNA linear EST 25-SEP-2001  
 DEFINITION 603056866p1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5206217 5', mRNA sequence.  
 ACCESSION B1766766  
 VERSION B1766766.1 GI:15758344  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 834)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
 Plate: L1AM1517 row: c column: 18  
 High quality sequence stop: 772.

FEATURES  
 source  
 1. 834  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5206217"  
 /clone\_lib="NIH\_MGC\_122"  
 /lab\_host="DH10B"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."

BASE COUNT 154 a 287 c 223 g 170 t  
 ORIGIN  
 Query Match 30.3%; Score 416; DB 13; Length 834;  
 Best Local Similarity 99.7%; Pred. No. 3.4e-203;  
 Matches 656; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 272 CCCAGAGAGCTGTGTGAGAGGAGCCAGGACCCGTCGGAATGATCCCGACAG 331  
 DB 1 CCCAGAGAGCTGTGTGAGAGGAGCCAGGACCCGTCGGAATGATCCCGACAG 60  
 OY 332 AAGAAAGCCAGATCTGCGCTTCTGAACGCACTAGTTGCGCTCGAGAACTGAC 391  
 DB 61 AAGAAAGCCAGATCTGCGCTTCTGAACGCACTAGTTGCGCTCGAGAACTGAC 120  
 OY 392 CTAAGGCCGGAACACAGGCTCGAAGAGCATGCGAGCCATTATGAAGTTATCCAC 451  
 DB 121 CTAAGGCCGGAACACAGGCTCGAAGAGCATGCGAGCCATTATGAAGTTATCCAC 180

QY	452	GACC - TTGACACAGGACGGAGGCGACAGCTGTGGACCGGACAGAGTATGTCTGGAGGAA	510
Db	181	GACGCTGGACAGGACGGAGGCGACAGCTGTGGACCGGACAGTATGTCTGGAGGAA	240
QY	511	GCCAGAAATCAACACGCTCCAGCCCTCTGGCTACACACCCGACAGATCGGGAGTTTATATGTC	570
Db	241	GCCAGAAATCAACACGCTCCAGCCCTCTGGCTACACACCCGACAGATCGGGAGTTTATATGTC	300
QY	571	ACCCGGGCTGGGGCTCTACTACCTGTACTGTACAGGTGCACCTTATATAGGGGAGAGGCTGTC	630
Db	301	ACCCGGGCTGGGGCTCTACTACCTGTACTGTACAGGTGCACCTTATATAGGGGAGAGGCTGTC	360
QY	631	TACCTGAAGCTGACCTTGCTGTGTGTGATGATGTGTGCTGGCCCTGGCGCTGCTCTGAGGAAATTC	690
Db	361	TACCTGAAGCTGACCTTGCTGTGTGTGATGATGTGTGCTGGCCCTGGCGCTGCTCTGAGGAAATTC	420
QY	691	TTACACCACTGGGGGCGCAATTCCTGTGGGGCCCAAGTCCGCTCTGTGCCAGATGTCTGGGGCT	750
Db	421	TTACACCACTGGGGGCGCAATTCCTGTGGGGCCCAAGTCCGCTCTGTGCCAGATGTCTGGGGCT	480
QY	751	TTGGGCCCTGGCGGCGCAGGGGTCCCTCCCTCGGATCGCACACCTCCCTGAGGCGCATATCCAG	810
Db	481	TTGGGCCCTGGCGGCGCA - GGTCTCTCCCTCGGATCCGACACCTCCCTGTGGCCATCTCCAG	539
QY	811	GCTGCCCCCTTCTCACCTTACTTGGACCTTTTCCAGGTTCACTGAGGGGCGCTGTCTCC	870
Db	540	GCTGCCCCCTTCTCACCTTACTTGGACCTTTTCCAGGTTCACTGAGGGGCGCTGTCTCC	599
QY	871	CCACAGTCGTCCAGAGGCTGGCGGCTGCCCTCGACACAGCTCTGTGGGACACCCGGTCCCT	928
Db	600	CCACAGTCGTCCAGAGGCTGGCGGCTGCCCTCGACACAGCTCTGTGTGGACACCCGGTCCCT	657

RESULT 11	
AI422796/c	
LOCUS	AI422796
DEFINITION	413 bp mRNA linear EST 30-MAR-1999
ACCESSION	U122606.1
VERSION	NCI-CGAP_Brn23 Homo sapiens CDNA clone IMAGE:2096962 3'
KEYWORDS	AI422796 AI422796 mRNA sequence.
SOURCE	AI422796.1 GI:4268727
ORGANISM	EST.
	human.
	Homo sapiens

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

1 (bases 1 to 413)  
NCI/NINDS-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CCGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.

Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: [www-bio.llni.gov/bbrp/image/image.html](http://www-bio.llni.gov/bbrp/image/image.html)  
Insert Length: 1728 Std Error: 0.00  
Seq primer: -400P from Gabc  
High quality sequence stop: 410.

```

FEATURES
source
location/Qualifiers
1. .413
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2096962"
/clone_1lb="NCI-GCAP_Brn23"
/russse_type="gliblastoma
/lab_host="DH10B"

```

BASE COUNT  
ORIGIN

100 a 88 c 119 g 106 t

/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 5' TGTACCAATCTGCAATGAGGAGCGCGGCATCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonalido."

Query Match	28.7%	Score 394	DB 9	Length 413
Best Local Similarity	100.0%	Pred. NO.	7.5e-192	
Matches 394	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy 980 GGTGGCCGTGGGGCTGTTACAGTGTTTTCATCCATCAATAAATACAGTATCCCACTCTTAT 1039  
Db 413 GGTGGCCGTGGGGCTGTTACAGTGTTTTCATCCATCAATAAATACAGTATCCCACTCTTAT 354

qy 1040 CTTACAACTCCCCCAGCGCCACTCTCCACTCACTAGCTCCCCAATCCCTGACCCCTTTG 1099  
 |||||  
 db 353 CTTACAACTCCCCCAGCGCCACTCTCCACTCACTAGCTCCCCAATCCCTGACCCCTTTG 294

**Oy** 1100 AGGCCCCCAGTGCATCTCGACTCCCCCTGGCCACAGACCCCAGGGCATTTGTTCAC TG 1155  
|||||  
**Dd** 293 AGGCCCCCA GTGATCTCGACTCCCCCTGGCCACAGACCCCAGGGCATTTGTTCAC TG 234

**Qy** 1160 TACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGCSCACTAAGAGGGGCTGCAC 1219  
|||||  
|||||  
**Db** 233 TACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGCSCACTTAAAGAGGGGCTGCAC 174  
|||||  
|||||

qy 1220 CTGGCGGCAAGAACCCAAAGAGACTGCGGCGTAGGCCAGGAGTTCCTCAATGTATAGGGCG 1219

Db 173 CTGGCGGCAAGAACCCAAAGAGACTGCGGCGTAGGCCAGGAGTTCCTCAATGTATAGGGCG 114

9y 1280 AGAAACAAGACAAGCTCTCCCTTGAGAAATCCCTGAGATTTTAAACAGATATTATT 13  
 |||||  
 Db 113 AGAAACAAGACAAGCTCTCCCTTGAGAAATCCCTGAGATTTTAAACAGATATTATT 54

Qy 1340 TTTATTATTATGTGACAAATGTTGTAATGG 13  
 |||||  
 Db 53 TTTATTATTATGTGACAAATGTTGTAATGG 20

RESULT	12
BM971606/c	
LOCUS	568 bp mRNA
DEFINITION	linear EST 21-MAR-2002
	UI-CF-EC1-abd-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens cdna clone
	UI-CF-EC1-abd-p-06-0-UI 3, mRNA sequence.

VERSION	BM971606.1	GI:19569193
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

1 (bases 1 to 568)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
57044477  
Contact: McCray, PB

University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: [paul-mccrory@uiowa.edu](mailto:paul-mccrory@uiowa.edu)  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed By: Dr. M. Bento Soares, University of Iowa

DNA Sequencing-by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA  
 sequence: 1-82, >AT-rich#low-complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLY-A-yes.

# FEATURES

## SOURCE

Location/Qualifiers

1..568

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-CF-EC1-ab1-p-06-0-UI"

/clone\_lib="UI-CF-EC1"

/tissue\_type="lung"

/dev\_stage="Adult and Fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a  
 modified polylinker; Site.1: EcoR I; Site.2: Not I;  
 UI-CF-EC1 is a normalized cDNA library containing the  
 following tissue(s): Normal lung from adult and from fetal  
 day 64, day 87, week 19 and week 42. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into p773-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 AAGTCTTAC.  
 TAG\_LIB-UI-CF-EC1  
 TAG\_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371  
 and 380-383  
 TAG\_SEQ=AACTGCTTAC"

## BASE COUNT

130 a 119 c 184 g 135 t

Query Match 28.7%; Score 394; DB 14; Length 568;  
 Best Local Similarity 100.0%; Pred. NO. 7.6e-192; Indels 0; Gaps 0;  
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

980 GCTGCTGGGCTGTCACGTTTTCATCCACATTAATACATATTCACCTCTTAT 1039  
 |||||||  
 421 GCTGCTGGGCTGTCACGTTTTCATCCACATTAATACATATTCACCTCTTAT 362  
 |||||||  
 1040 CTACACATCCCGCCACCGCCACCTCTCACTAGCTCCCAATCCCTGACCTTGG 1099  
 |||||||  
 361 CTACACATCCCGCCACCGCCACCTCTCACTAGCTCCCAATCCCTGACCTTGG 302  
 |||||||  
 1100 AGGCCCCAGATGTCAGTCCCGCCGAGACAGACCCCGAGGCGATGTCACAG 1159  
 |||||||  
 301 AGGCCCCAGATGTCAGTCCCGCCGAGACAGACCCCGAGGCGATGTCACAG 242  
 |||||||  
 1160 TACTCTGTGGGCAAGATGGTTCAGAAAGCCCACTTCAAGCACTAAGAGGGGCTGAC 1219  
 |||||||  
 241 TACTCTGTGGGCAAGATGGTTCAGAAAGCCCACTTCAAGCACTAAGAGGGGCTGAC 182  
 |||||||  
 1220 CTGGGGGAGAGAGCCAAAGAGACTGGGCTTAGGCCAGAGAGTTCCTCAATGTGAGGGGCG 1279  
 |||||||  
 181 CTGGGGGAGAGAGCCAAAGAGACTGGGCTTAGGCCAGAGAGTTCCTCAATGTGAGGGGCG 122  
 |||||||  
 1280 AGAAACAGACAGCTCTCCCTTGAAGATTCCTGATTTTAAACAGATATTTAT 1339  
 |||||||  
 121 AGAAACAGACAGCTCTCCCTTGAAGATTCCTGATTTTAAACAGATATTTAT 62  
 |||||||  
 1340 TTTATTTATTTGACAAAATGTGATTAATGG 1373  
 |||||||  
 61 TTTATTTATTTGACAAAATGTGATTAATGG 28  
 |||||||

RESULT 13  
 B1824443

LOCUS B1824443 531 bp mRNA linear EST 04-OCT-2001

DEFINITION 603038693P1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179510 5',

ACCESSION B1824443

VERSION B1824443.1 GI:15935993

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mhc.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cga@bbs-remail.nih.gov

source Tissue Procurement: Life Technologies, Inc.

Location/Qualifiers

1..531

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5179510"

/clone\_lib="NIH\_MGC\_115"

/lab\_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:  
 pCMV-Sport6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA  
 source anonymous pool of 6 male brains, age range 23-27; 1  
 male lung, age 27; and 1 male testis, age 69. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.8 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 021. Note: this is a NIH-MGC Library."

## BASE COUNT

72 a 223 c 119 g 117 t

Query Match 27.3%; Score 375; DB 13; Length 531;  
 Best Local Similarity 99.8%; Pred. No. 4.8e-182; Indels 0; Gaps 0;  
 Matches 425; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

767 GGTGCTCCCTGGGAGTCCGACCCCTGCGGCGCATCTCAAGGCTGCCCCCTTCCTCA 826  
 |||||||  
 106 GGTGCTCCCTGGGAGTCCGACCCCTGCGGCGCATCTCAAGGCTGCCCCCTTCCTCA 165  
 |||||||  
 827 CCTACTTGGGAGTCTTTCAGAGTTCAGAGGAGGAGGAGTCTCCACAGTGTCCAGG 886  
 |||||||  
 166 CCTACTTGGGAGTCTTTCAGAGTTCAGAGGAGGAGGAGTCTCCACAGTGTCCAGG 225  
 |||||||  
 887 CTGGGGGCTCCCTTGCAGAGCTCTGTGGGACCCGAGTCCCTGTGCCCCACCTTAGCCG 946  
 |||||||  
 226 CTGGGGGCTCCCTTGCAGAGCTCTGTGGGACCCGAGTCCCTGTGCCCCACCTTAGCCG 285  
 |||||||  
 947 CTCTTGTCCAGAGCTGCCCCCTGCTAGAGGAGTGGGCTGGGCGCTGTTACAGTCTTTC 1006  
 |||||||  
 286 CTCTTGTCCAGAGCTGCCCCCTGCTAGAGGAGTGGGCTGGGCGCTGTTACAGTCTTTC 345  
 |||||||  
 1007 CATCCACATAAATACAGATATCCCACTTATTTTACAAATCCCCCAACCGCCCACTCTC 1066  
 |||||||  
 346 CATCCACATAAATACAGATATCCCACTTATTTTACAAATCCCCCAACCGCCCACTCTC 405  
 |||||||  
 1067 CACCTCAGTACGTCCTCCCAATCCCTAGACCTTTGAGGCCCCGAGTATCTGACTCCCCC 1126  
 |||||||  
 406 CACCTCAGTACGTCCTCCCAATCCCTAGACCTTTGAGGCCCCGAGTATCTGACTCCCCC 465  
 |||||||





BASE COUNT	ORIGIN
165 a	synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIM-MCC Library."
286 c	315 g 184 t 1 others

BASE COUNT	165 a	286 c	315 g	184 t
ORIGIN				

ORIGIN

Query Match	25.48;	Score 349;	DB 14;	Length 951;
Best local similarity	100.0%;	Prod No	1.3e-168;	

Best Local Similarity	100.0%;	Pred. No. 1.3e-168;
Matches 349;	Conservative 0;	Mismatches 0;
		Indels 0;
		Gaps 0;

Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	369	AGTTGGCGCTGCGCAGAGTGTGACCTTAAGGCGGAAAAACAGCGGCTCGAAGACGATCGC	428
Db	1	AGTTGGCGCTGCGCAGAGTGTGACCTTAAGGCGGAAAAACAGCGGCTCGAAGACGATCGC	60
QY	429	AGCCCATTTATTAATTCATCCAGACCTGGACAGACGAGACGACAGCTGTGACGG	488
Db	61	AGCCCATTTATTAATTTATCCACAGCCTGGACAGAGCGAGACGACAGCTGTGACCG	120
QY	489	GACAGTGTAGTGTGCGGAGGAGACAGATATACAGCTCCAGCCTTGTGCTACAACCG	548
Db	121	GACAGTGTAGTGTGCGGAGGAGACAGATATACAGCTCCAGCCTTGTGCTACAACCG	180
QY	549	CCATATGGGGAGTTTATAGTACACCGCGGCTGTACTACTGTACATCTCAGTGC	608
Db	181	CCATATGGGGAGTTTATAGTACACCGCGGCGTGTACTACTGTACATCTCAGTGC	240
QY	609	CTTTGATGAGGGGAAAGCGCTGTACCTCGAAGCTGAGCTTGCTGGTGAATGCTGTGCG	668
Db	241	CTTTGATGAGGGGAAAGCGCTGTACCTCGAAGCTGAGCTTGCTGGTGAATGCTGTGCG	300
QY	669	CCCTGGGCTCGCCTGGAGGAATTCATAGGCATCGCGGGCCAGTTCCCTGCGG	717
Db	301	CCCTGGGCTCGCCTGGAGGAATTCATAGGCATCGCGGGCCAGTTCCCTGCGG	349

Search completed: May 8, 2003, 13:50:17  
Job time : 1862 secs

**THIS PAGE BLANK (uspro)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 10:15:30 ; Search time 342 Seconds  
(without alignments)  
9040.917 Million cell updates/sec

Title: US-09-245-198a-3  
Perfect score: 1373  
Sequence: 1 atgcatgttgtagcttga.....gacaaatgtgtaaatg 1373

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N.GeneSeq\_101002.\*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1373	100.0	1373	19	AAV18600
2	1247	90.8	1364	24	ABK34881
3	1172	85.4	1353	21	AAA49717
4	1172	85.4	1353	24	ABK40255
5	1172	85.4	1421	20	AAK56000
6	958	69.8	1236	19	AAV47613
7	958	69.8	1236	22	AAD04350
8	784	57.1	1030	20	AAK23424
9	625	45.5	898	22	AAK03964

10	179	13.0	195	24	ABK29540	Colon adenocarcino
11	76	5.5	282	16	AAT22190	Human gene signatu
12	60	4.4	60	24	ABN41049	Human spliced tran
13	60	4.4	60	24	ABN58591	Human spliced tran
14	60	4.4	60	24	ABN58593	Human spliced tran
15	60	4.4	60	24	ABN58848	Human spliced tran
16	60	4.4	60	24	ABN58848	Human spliced tran
17	50	3.6	50	20	AAK56002	Human tumour necro
18	50	3.6	50	21	AAA49733	Human P0207 DNA p
19	50	3.6	50	24	ABK40292	Oligonucleotide pr
20	46	3.4	701	20	AAK23425	Mouse TNFRL3 DNA.
21	22	3.4	1168	19	AAV18599	Mus musculus tumou
22	26	1.9	26	24	ABK40356	Reverse PCR primer
23	26	1.9	147	16	AAT19717	Human gene signatu
24	26	1.9	153	24	AAT69032	Activated T-cell d
25	26	1.9	281	22	AAS24709	Human ovarian PCR-
26	26	1.9	386	24	ABQ60530	Human colon cancer
27	26	1.9	391	22	AAH83337	Human ovarian tumo
28	26	1.9	439	24	ABN96246	Gene #2744 used to
29	26	1.9	439	24	ABL62690	Colon adenocarcino
30	26	1.9	452	21	AAC06922	Human secreted pro
31	26	1.9	483	24	ABK55088	Human colon cancer
32	26	1.9	486	21	AAZ51563	Human hypoxia resp
33	26	1.9	487	26	ABN94231	Gene #729 used to
34	26	1.9	516	21	AAC01272	Human secreted pro
35	26	1.9	531	26	ABK55407	Human colon cancer
36	26	1.9	540	21	AAC01271	Human secreted pro
37	26	1.9	570	21	AAC77898	Human cancer assoc
38	26	1.9	580	21	AAZ51562	Human hypoxia resp
39	25	1.8	400	24	ABQ58318	Human colon cancer
40	25	1.8	626	23	AAK56426	DNA encoding novel
41	24	1.7	24	20	AAK23450	Human TNFRL3 RACE p
42	24	1.7	24	20	AAK23451	Human TNFRL3 RACE p
43	24	1.7	40	20	AAK56003	Human tumour necro
44	23	1.7	23	24	ABK40355	Probe for gene amp
45	22	1.6	38	20	AAK56004	Human tumour necro

## ALIGNMENTS

RESULT 1	
AAV18600	
ID	AAV18600 standard; cDNA; 1373 BP.
AC	AAV18600; .
XX	
DT	21-JUL-1998 (first entry)
XX	
DE	Homo sapiens tumour necrosis factor related ligand (TNFRL) gene.
XX	
TRELL	tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection; ds.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	1..852
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
XX	
PN	WO9805783-A1.
XX	
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
PR	18-OCT-1996; 96US-0028515.
XX	
PA	(BIOT ) BIOGEN INC.

PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicheportliche Y;

XX WPI: 1998-145619/13.

DR P-PSDB; AAM47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts

PS Claim 2: Pages 48-50; 69pp; English.

XX The sequence is that encoding human tumour necrosis factor related  
CC ligand (TREML). TREML or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TREML  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TREML-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon-gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TREML and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TREML-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TREML.  
CC It may also be of use in the preparation of probe probes for  
CC screening natural/synthetic DNAs for TREML-encoding sequences  
CC and for antisense therapy.

SO Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:

Query Match 100.0%; Score 1373; DB 19; Length 1373;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATGTAGTAACTTGAATTTCCGCGCGGCTCCCTCCCGCATCTCTGGG 60  
DB 1 ATGTCATGTAGTAACTTGAATTTCCGCGCGGCTCCCTCCCGCATCTCTGGG 60  
QY 61 TCCCGGATGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
DB 61 TCCCGGATGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
QY 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 181 CTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
DB 181 CTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
QY 241 AGCGGGGATGCTGTCCGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
DB 241 AGCGGGGATGCTGTCCGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
QY 301 CAGAGCCGCTCGAATGAAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 301 CAGAGCCGCTCGAATGAAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 361 AACCGACTAGTTCGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
DB 361 AACCGACTAGTTCGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
QY 421 GCGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
DB 421 GCGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 481 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
DB 481 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 541 TACAAACCGGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
DB 541 TACAAACCGGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
QY 601 CAGGTGACATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
DB 601 CAGGTGACATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
QY 661 GTGCTGGGCGGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
DB 661 GTGCTGGGCGGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
QY 721 CAGCTCCGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
DB 721 CAGCTCCGCTTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
QY 781 ATCCGACACCTTCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
DB 781 ATCCGACACCTTCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
QY 841 TTCCAGGTTCACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 841 TTCCAGGTTCACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 901 CGACAGCTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
DB 901 CGACAGCTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
QY 961 CCGTGGGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
DB 961 CCGTGGGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
QY 1021 ACAGATATCCCACTTATTTTACAACTCCCAAGGAGGAGGAGGAGGAGGAGGAG 1080  
DB 1021 ACAGATATCCCACTTATTTTACAACTCCCAAGGAGGAGGAGGAGGAGGAGGAG 1080  
QY 1081 CCCAATCCCTGACCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
DB 1081 CCCAATCCCTGACCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1141 CAGGCGATGTTGACATGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 1141 CAGGCGATGTTGACATGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1201 GCACATAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
DB 1201 GCACATAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
QY 1261 TTCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
DB 1261 TTCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
QY 1321 TTTTAAACAGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 1373  
DB 1321 TTTTAAACAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373

RESULT 2  
ABK34881  
ID ABK34881 standard; cDNA; 1364 BP.  
XX  
AC ABK34881;  
DT 08-MAY-2002 (first entry)  
XX  
DE Human cDNA encoding secreted protein #19.  
XX  
KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KM coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
 KW tissue regeneration; wound healing; burn; hematopoiesis;  
 KM myeloid cell deficiency; lymphoid cell deficiency.  
 XX Homo sapiens.  
 OS  
 PN WO200177288-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 29-MAR-2001; 2001WO-US10224.  
 XX  
 PR 06-APR-2000; 2000US-195582P.  
 XX  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Wong GC, Clark HF, Pechtel K, Agostino MJ, Howes SH, Resnick RJ;  
 PI Gulukota K, Graham JR;  
 XX  
 DR WPI; 2002-179321/23.  
 XX  
 PT Five hundred and ninety two polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders  
 PS  
 XX  
 PS Claim 1; Page 82; 372pp; English.  
 CC The invention relates to 592 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins. The polynucleotides can be used as probes for the  
 CC identification and isolation of full length cDNA and genomic DNA. The  
 CC polynucleotides and proteins can also be used as nutritional supplements.  
 CC The proteins are useful in the treatment of various immune deficiencies  
 CC and disorders such as viral infections, bacterial infections, fungal  
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
 CC and conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment of  
 CC burns, infections and ulcers. The proteins are also useful for regulating  
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
 CC Sequences ABK34863-ABK3454 represent polynucleotides of the invention.  
 CC  
 XX  
 SQ Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;  
 Query Match 90.8%; Score 1247; DB 24; Length 1364;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1347; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 305 CAGACGAAAGAAAGCCAGATCTCGGCTTCTCTAACCAGTACGTTGGCTGCACAA 364  
 QY 385 AGTGACCTAAAGCCGGAAACACGGGCTCGAAGAGCATGCAGCCCATTAATGATT 444  
 Db 365 AGTGACCTAAAGCCGGAAACACGGGCTCGAAGAGCATGCAGCCCATTAATGATT 424  
 QY 445 CATCCAGACCTGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 504  
 Db 425 CATCCAGACCTGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 484  
 QY 505 GAGGAAGCCAGATCAACAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 564  
 Db 485 GAGGAAGCCAGATCAACAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 544  
 QY 565 ATAGTCACCCGGGCTGGGCTTACTACCTGTAAGTGTGAGTGTGAGTGTGAGTGT 624  
 Db 545 ATAGTCACCCGGGCTGGGCTTACTACCTGTAAGTGTGAGTGTGAGTGTGAGTGT 604  
 QY 625 GGTGTACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 684  
 Db 605 GGTGTACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 664  
 QY 685 GAATTCACGACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 744  
 Db 665 GAATTCACGACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 724  
 QY 745 GGGCTTTGGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 804  
 Db 725 GGGCTTTGGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784  
 QY 805 CTCAGAGCTGACCCCTTCTCTACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 864  
 Db 785 CTCAGAGCTGACCCCTTCTCTACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 844  
 QY 865 GTCCTCCACAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 924  
 Db 845 GTCCTCCACAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 904  
 QY 925 CCTCTGCCCCACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 984  
 Db 905 CCTCTGCCCCACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964  
 QY 985 CTGGGCTGTACAGCTGTTTCCATCCACATTAATGAGTATTCCTGATCTTAC 1044  
 Db 965 CTGGGCTGTACAGCTGTTTCCATCCACATTAATGAGTATTCCTGATCTTAC 1024  
 QY 1045 AACTCCCCACGCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104  
 Db 1025 AACTCCCCACGCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1084  
 QY 1105 CCCAGTATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1164  
 Db 1085 CCCAGTATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144  
 QY 1165 TGTGGCAAGATGGTCCAGAAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1224  
 Db 1145 TGTGGCAAGATGGTCCAGAAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1204  
 QY 1225 GGCAGAGCCAAAGAGTGGGCTGAGCCAGAGTTCCTCAATGTGAGGGGCGAGAA 1284  
 Db 1205 GGCAGAGCCAAAGAGTGGGCTGAGCCAGAGTTCCTCAATGTGAGGGGCGAGAA 1264  
 QY 1285 CAAGACAAGCTCTCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTGAGAT 1344  
 Db 1265 CAAGACAAGCTCTCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTGAGAT 1324  
 QY 1345 TATTATTGACAAATGTTGATTAATGG 1373  
 Db 1325 TATTATTGACAAATGTTGATTAATGG 1353

RESULT 3

AA	AAA49717	
ID	AAA49717 standard; cDNA; 1353 BP.	
XX		
AC	AAA49717;	
DT	25-SEP-2000 (first entry)	
XX		
DE	Human PRO207 cDNA clone DNA30879-1152.	
XX		
KW	PRO207; human; antitumour; tumour; therapy; cytostatic; breast cancer; ovarian cancer; renal cancer; colorectal cancer; uterine cancer; prostate cancer; lung cancer; bladder cancer; central nervous system cancer; melanoma; leukaemia; neoplasm; ss.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	58..807
FT		/*tag= a
FT	sig_peptide	58..177
FT		/*tag= b
FT	mat_peptide	178..804
FT		/tag= c
XX		
PN	MO20037638-A2.	
XX		
PD	29-JUN-2000.	
XX		
PE	02-DEC-1999;	99MO-US28565.
XX		
PR	22-DEC-1998;	98US-0113296.
PR	08-MAR-1999;	99MO-US05028.
PR	21-APR-1999;	99US-0130232.
PR	28-APR-1999;	99US-0131445.
PR	14-MAY-1999;	99US-0134287.
PR	20-JUL-1999;	99US-0144758.
PR	26-JUL-1999;	99US-0145698.
PR	15-SEP-1999;	99MO-US21090.
XX	15-SEP-1999;	99MO-US21547.
PA	(GENE ) GENENTECH INC.	
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA; Napier MA, Pittl RM, Wood WI;	
XX		
DR	WPI: 2000-442668/38.	
DR	P-PDB: AA95338.	
PT		
PT	Novel composition to inhibit neoplastic cell growth or for treating tumour in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO321, PRO324, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866	
XX		
PS	Claim 20; Fig 3; 172pp: English.	
XX		
CC	The present sequence is that of cDNA clone DNA30879-1152 (ATCC 209358) encoding human PRO207 (see AA95338), which shows homology to several members of the tumour necrosis factor family, especially human lymphotoxin (23.4%). The cDNA was identified in a foetal kidney cDNA library following identification of an expressed sequence tag with homology to human Apo-2 ligand. A claimed method for inhibiting the growth of a tumour cell comprises exposing the tumour cell to PRO179, PRO207, PRO320, PRO321, PRO324, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see AA95337-49), their agonists or chimeric polypeptides incorporating them. The tumour is especially a cancer selected from breast, ovarian, renal, colorectal, uterine, prostate, lung, bladder and central nervous system cancer, melanoma and leukaemia. Nucleic acids encoding PRO179 etc. are used in the recombinant production of the antitumour polypeptides.	
XX		
QD	Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:	

Query Match		85.48	Score 1172	DB 21	Length 1353
Best Locality		Similarity 99.88	Pred. No.		
Matches 1322		Conservative 0	Mismatches 3	Indels 0	Gaps
QY	CGATCCTTCGGGCTCCCGGGATGGGGGGCGGTGAGGACAGGACGCCGCCGCCCATG	108			
QY	1	CGATCCTTCGGGCTCCCGGGATGGGGGGCGGTGAGGACAGGACGCCGCCGCCCATG	60		
QY	109	GGCGGCGGTGCGAGCCAGAGGCGGAGGGGGGGCGCGGGGGAGGCGCGGGGCAACCCCTGCTG	168		
DB	61	GGCGGCGGTGCGAGCCAGAGGCGGAGGGGGGGCGCGGGGGAGGCGCGGGGCAACCCCTGCTG	120		
QY	169	GTCCCGCTCGGCGCTGGGCGCTGGGCGCTGGCGCTGGCGCTCGGCGCTCTGCTGGCGCTG	228		
DB	121	GTCCCGCTCGGCGCTGGGCGCTGGGCGCTGGCGCTGGCGCTGGCGCTCTGCTGGCGCTG	180		
QY	229	GTCAGTTTGGGGAGCCGGGCACTGCTCTCCGCCAGAGACCTCTCCAGAGAGAGTGGTG	288		
DB	181	GTCAATTTTGGGGAGCCGGGCACTGCTCTCCGCCAGAGAGCTCTCCAGAGAGAGTGGTG	240		
QY	289	GCAGAGAGAGACCAAGACCCCTGGGAATCAATCCCGACAGAGAAAGCCAGATCTT	348		
DB	241	GCAGAGAGAGACCAAGACCCCTGGGAATCAATCCCGACAGAGAAAGCCAGATCTT	300		
QY	349	GGCGCTTTCCTGAACCGACTAGTTGCGGCTGCGAGAGATGACACTAAAGGCCGGAATA	408		
DB	301	GGCGCTTTCCTGAACCGACTAGTTGCGGCTGCGAGAGATGACACTAAAGGCCGGAATA	360		
QY	409	CGGGCTCGAAGAGAGATCGGAGCCCATTAATGAAGTTCAATCAAGACCTGGACAGAGCA	468		
DB	361	CGGGCTCGAAGAGAGATCGGAGCCCATTAATGAAGTTCAATCAAGACCTGGACAGAGCA	420		
QY	469	GGCGAGGACAGGTGTGAGACGGGACAGTGAAGTGGGTGGAGAGAAAGCCAGATCAACAGCTCC	528		
DB	421	GGCGAGGACAGGTGTGAGACGGGACAGTGAAGTGGGTGGAGAGAAAGCCAGATCAACAGCTCC	480		
QY	529	AGCCCTTCGGGCTTCAACACCGCAGATGGGGAGTTTATAGTCAACCCGGGCTGGGCTTAC	588		
DB	481	AGCCCTTCGGGCTTCAACACCGCAGATGGGGAGTTTATAGTCAACCCGGGCTGGGCTTAC	540		
QY	589	TACCTGTACTGTCAAGGTGCATTTTGATGAGGGGAAAGGCTGTCACTGAAGCTGAGACTTG	648		
DB	541	TACCTGTACTGTCAAGGTGCATTTTGATGAGGGGAAAGGCTGTCACTGAAGCTGAGACTTG	600		
QY	649	CTGTGTGATGTGTGTGTGGCCCTGCGCTGCTGAGAGAAATTTCAAGCACTCGCGCACT	708		
DB	601	CTGTGTGATGTGTGTGTGGCCCTGCGCTGCTGAGAGAAATTTCAAGCACTCGCGCACT	660		
QY	709	TCCCTGCGGCGCCAGCTTCGCGCTGTGCGCAAGTGTGAGGCTGTGTGGCCCTGCGCGCAGAG	768		
DB	661	TCCCTGCGGCGCCAGCTTCGCGCTGTGCGCAAGTGTGAGGCTGTGTGGCCCTGCGCGCAGAG	720		
QY	769	TCTTCCTGCGGATCGGCAACCTTCCTCCCTGGGCCAATCTCAAGGCTGCCCCCTTCTCAAC	828		
DB	721	TCTTCCTGCGGATCGGCAACCTTCCTCCCTGGGCCAATCTCAAGGCTGCCCCCTTCTCAAC	780		
QY	829	TACTTGGAGACTTTCAGAGTTCACTGAGGGGGCCCTGCTGCCCAAGTCGTCGCCAGGCT	888		
DB	781	TACTTGGAGACTTTCAGAGTTCACTGAGGGGGCCCTGCTGCCCAAGTCGTCGCCAGGCT	840		
QY	889	GGCGGCTCCCTCGACAGCTCTGCGGCAACCCCGTCCCTGTGCCCCACCCCTCAACCGCT	948		
DB	841	GGCGGCTCCCTCGACAGCTCTGCGGCAACCCCGTCCCTGTGCCCCACCCCTCAACCGCT	900		
QY	949	CTTTGTCTCAGACCTGCGCCCTCCCTCTAAGAGGTGCTCTGGGCTGTTCACGTGTTTCA	1008		
DB	901	CTTTGTCTCAGACCTGCGCCCTCCCTCTAAGAGGTGCTCTGGGCTGTTCACGTGTTTCA	960		
QY	1009	TCCCACTAAATACAGATATCCACACTTATCTTAACTCCGCCACCGCCACACTCTCA	1068		
DB	961	TCCCACTAAATACAGATATCCACACTTATCTTAACTCCGCCACCGCCACACTCTCTCA	1020		
QY	1069	CCTCAGTAGTCCCCCAATCCCTGACCCCTTTGAAGGCCCAAGATCTCGACTCCCCCTCG	1128		





Db	721	TCCTCCGTGGGGAATCCGACACCTTCCCTGGGCCAATCTCAAGAGCTGCCCTTCTCTCAC	780
Qy	829	TACTTTGGACTCTTCCAGGTTCACTAGAGGGCCCTGCTCTCCCAAGTCTGCCAGCT	888
Db	781	TACTTTGGACTCTTCCAGGTTCACTAGAGGGCCCTGCTCTCCCAAGTCTGCCAGCT	840
Qy	889	GCGGCGTCCCTCGACAGCTCTGTGGGCAACCGGTCCTCCCTGCCCAACCTCAGCGCT	948
Db	841	GCGGCGTCCCTCGACAGCTCTGTGGGCAACCGGTCCTCCCTGCCCAACCTCAGCGCT	900
Qy	949	CTTTGTCTCAGACCTGCCCTCCCTCTAGAAGCTGCCTGGGCTGTTCACTGTTTTC	1008
Db	901	CTTTGTCTCAGACCTGCCCTCCCTCTAGAAGCTGCCTGGGCTGTTCACTGTTTTC	960
Qy	1009	TCCGACATTAATACAGTATTCGACACCTTATCTTAAATCCCCACCGCCACACTCTCA	1068
Db	961	TCCGACATTAATACAGTATTCGACACCTTATCTTAAATCCCCACCGCCACACTCTCA	1020
Qy	1069	CCTGACTAGTCCCCCAATCCCTGACCCCTTTGAGGCCCCAGTAGATCTGCATCCCCCTG	1128
Db	1021	CCTGACTAGTCCCCCAATCCCTGACCCCTTTGAGGCCCCAGTAGATCTGCATCCCCCTG	1080
Qy	1129	GCCACAGACCCCAAGGGCACTGTGTCTACTGTACTCTGTGGGCAAGATGGGTCCAGAG	1188
Db	1081	GCCACAGACCCCAAGGGCACTGTGTCTACTGTACTCTGTGGGCAAGATGGGTCCAGAG	1140
Qy	1189	ACCCCACTTAGGCACTAAGAGGGGCTGACCTGCGCGAGGAGAACCAAGAGACTGGGC	1248
Db	1141	ACCCCACTTAGGCACTAAGAGGGGCTGACCTGCGCGAGGAGAACCAAGAGACTGGGC	1200
Qy	1249	CTAGGCAAGAGTCCCAATGTGAGGGCGGAGAAACAAGACAACTCCTCCTTGAGAA	1308
Db	1201	CTAGGCAAGAGTCCCAATGTGAGGGCGGAGAAACAAGACAACTCCTCCTTGAGAA	1260
Qy	1309	TTTCCCTGTGATTTTAAACAGATATTATTTTATTTATTTATTTGTGACAAATGTTGATA	1368
Db	1261	TTTCCCTGTGATTTTAAACAGATATTATTTTATTTATTTATTTGTGACAAATGTTGATA	1320
Qy	1369	AATGG 1373	
Db	1321	AATGG 1325	

```

Db 755 TCCTCCCTGCGATCGGACCCCTCCCTGGCCCAATCTCAAGCTGCCCCCTTCCTCACC 814
QY 828 TACTTCGAGCTCTTCAGAGTTCTACGTAGAGGGCCCTGGTCTCCCAAGTCTGCCAGGCT 888
Db 815 TACTTCGAGCTCTTCAGAGTTCTACGTAGAGGGCCCTGGTCTCCCAAGTCTGCCAGGCT 874
QY 889 GCGGCTCCCTGAGAGCTCTCTGGGCAACCGGTCCTCCCTGGCCCACTCCAGCCGCT 948
Db 875 GCGGCTCCCTGAGAGCTCTCTGGGCAACCGGTCCTCCCTGGCCCACTCCAGCCGCT 934
QY 949 CTTTCCTCAGAGCTGCCCCCTCCCTTAGAGCTGCTGGCCCTGTTCACTGTTTCCA 1008
Db 935 CTTTCCTCAGAGCTGCCCCCTCCCTTAGAGCTGCTGGCCCTGTTCACTGTTTCCA 994
QY 1009 TCCCACTAATATACAGTATCTCCACTCTATCTTACAACTCCCCCAAGCCCACTCTCA 1068
Db 995 TCCCACTAATATACAGTATCTCCACTCTATCTTACAACTCCCCCAAGCCCACTCTCA 1054
QY 1069 CCTCACTAGCTCCCAATCCCTGACCTTTAGAGCCCCCAAGATCTGACTCCCCCTG 1128
Db 1055 CCTCACTAGCTCCCAATCCCTGACCTTTAGAGCCCCCAAGATCTGACTCCCCCTG 1114
QY 1129 GCGACAGACCCCAAGGCAATCTGTCTCACTGACTCTGTGGCAAGATGGTCCAGAA 1188
Db 1115 GCGACAGACCCCAAGGCAATCTGTCTCACTGACTCTGTGGCAAGATGGTCCAGAA 1174
QY 1189 ACCCACTTCAAGCACTAGAGGGGCTGACCTGGGCGGAGAGCAAGCAAGAGCTGGCC 1248
Db 1175 ACCCACTTCAAGCACTAGAGGGGCTGACCTGGGCGGAGAGCAAGCAAGAGCTGGCC 1234
QY 1249 CTAGCCAGAGATTCCTCAATGTGAGAGGGCGAGAAACAAGACAGCTCTCCCTGAGAA 1308
Db 1235 CTAGCCAGAGATTCCTCAATGTGAGAGGGCGAGAAACAAGACAGCTCTCCCTGAGAA 1294
QY 1309 TTTCCCTGTGAGATTTTAAACAAGATATTTTATTTATTTATTTATTTATTTATTTATTT 1368
Db 1295 TTTCCCTGTGAGATTTTAAACAAGATATTTTATTTATTTATTTATTTATTTATTTATTT 1354
QY 1369 AATGG 1373
Db 1355 AATGG 1359

```

RESULT 6  
AAV47613  
ID AAV47613 standard; CDNA; 1236 BP.

```

AC AAV47613;
DE 27-OCT-1998 (first entry)
DE TNF related endothelium proliferative agent gene.
DE ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
DE tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..750
XX FT /*tag- a
XX FT /product= "TREPA"
XX
XX PM W09835061-A2.
XX PD 13-AUG-1998.
XX PF 12-FEB-1998; 98WO-US02859.
XX PR 10-FEB-1998; 98US-0021706.
XX PR 12-FEB-1997; 97US-0798692.
XX

```

(ABRO ) ABBOTT LAB.  
PI Wiley SR;  
DR WPI: 1998-447255/38.  
DR P-PSDB: AAM29745.  
XX  
XX Detecting nucleic acid encoding TREPA - useful for diagnosis and  
XX treatment of autoimmune disease, tumours and inflammation  
XX  
XX Claim 11; Page 123-4; 142pp; English.

The TNF-related endothelium proliferative agent (TREPA), or its  
activators or agonists, are used to treat a deficit of TREPA, e.g. to  
promote wound healing or tissue grafting, by promoting vascularisation,  
also to induce apoptosis for treating cancer and eliminating autoreactive  
T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
TREPA peptides can also be used to target cytotoxic agents or for  
affinity isolation of the corresponding receptor, the nucleic acid for  
which can be used to transform tumour cells to render them more  
responsive to TREPA and to screen for TREPA mimics.  
Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
vascularisation), inflammation or a wide range of autoimmune conditions,  
conditions involving abnormal stimulation of epithelial cells (e.g.  
atherosclerosis), for birth control (inhibiting ovulation and placental  
formation) or other angiogenic conditions (e.g. ulcers).

Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Query Match 69.8%; Score 958; DB 19; Length 1236;

Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1208; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 129 GCGAGAGGGGCGCGGGGGGAGCCCGGACCGCCCTGCTGTCGCCCTCGCGGGCT 188
Db 24 GCGAGAGGGGCGCGGGGGGAGCCCGGACCGCCCTGCTGTCGCCCTCGCGGGCT 83
QY 189 GGGCGTGGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 248
Db 84 GGGCGTGGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 143
QY 249 ATCGCTGTCCGCCAGAGAGCTGCTCCAGAGAGAGTGTGGAGAGAGAGAGAGAGAGAGAG 308
Db 144 ATCGCTGTCCGCCAGAGAGCTGCTCCAGAGAGAGTGTGGAGAGAGAGAGAGAGAGAGAG 203
QY 309 GTTCGAACTGAATCCCGACAGAAAGAAAGCAGAGATCCTGCGCTTCTCGAACCGACT 368
Db 204 GTTCGAACTGAATCCCGACAGAAAGAAAGCAGAGATCCTGCGCTTCTCGAACCGACT 263
QY 369 AGTTGGGCTCGCGAGAGAGTGCACCTTAAAGGCGGAGAAACAGGGGCTCGAAGAGCATGGC 428
Db 264 AGTTGGGCTCGCGAGAGAGTGCACCTTAAAGGCGGAGAAACAGGGGCTCGAAGAGCATGGC 323
QY 429 AGCCCAATTAATGAATTCATCCAGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 488
Db 324 AGCCCAATTAATGAATTCATCCAGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 383
QY 489 GACAGTGAAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
Db 384 GACAGTGAAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443
QY 549 CCAGATCGGGGAGTTTATATGATCAGCCGGGCTGCTACTACTGCTACTGCTACTGCTACT 608
Db 444 CCAGATCGGGGAGTTTATATGATCAGCCGGGCTGCTACTACTGCTACTGCTACTGCTACT 503
QY 609 CTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
Db 504 CTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 563
QY 669 CCTGCGCTGCTCGAGAGAAATTCAGAGCACTGCGGCGCAATTCCTCGGCGCCCACTCGG 728
Db 564 CCTGCGCTGCTCGAGAGAAATTCAGAGCACTGCGGCGAGATTCCTCGGCGCCCACTCGG 623

```



```
OY 729 CCTCTGACAGGTGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 788
DB 624 CCTCTGACAGGTGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 683
OY 789 CCTCTGACAGGTGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 848
DB 684 CCTCTGACAGGTGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 743
OY 849 TCACGAGAGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 908
DB 744 TCACGAGAGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 803
OY 909 CTCTGAGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 968
DB 804 CTCTGAGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 863
OY 969 TCCTCTGAGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1028
DB 864 TCCTCTGAGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 923
OY 1029 CCCACTCTATCTTACACTCTCCCAACCGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1088
DB 924 CCCACTCTATCTTACACTCTCCCAACCGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 983
OY 1089 CTGACCTCTTGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1148
DB 984 CTGACCTCTTGGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1043
OY 1149 TGTGTTCACTGCTGCTGTGTGGGCAAGAGTGGTTCAGAAAGCCCACTTCAGGCACTAAG 1208
DB 1044 TGTGTTCACTGCTGCTGTGTGGGCAAGAGTGGTTCAGAAAGCCCACTTCAGGCACTAAG 1103
OY 1209 AGGGCTGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1268
DB 1104 AGGGCTGAGGCTGAGGCTGTTGGCCCTGCGCCAGGAGTCTCCCTGCGAGTCCGCAC 1163
OY 1269 TGTGAGGAGGCGAGAAACAGACAGCTCTCCCTGAGAAATTCCTGTGTGATTTTAAAA 1328
DB 1164 TGTGAGGAGGCGAGAAACAGACAGCTCTCCCTGAGAAATTCCTGTGTGATTTTAAAA 1223
OY 1329 CAGATATTTATTTT 1341
DB 1224 CAGATATTTATTTT 1236

RESULT 8
ID AAX23424
AAX23424 standard; DNA; 1030 BP.
AC AAX23424;
XX 18-JUN-1999 (first entry)
DE Human TNRL3 DNA.
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
XX developmental abnormality; gestational abnormality; prostate cancer;
XX APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX apoptosis; human; ss.
OS Homo sapiens.
XX Key 1..627 Location/Qualifiers
XX CDS /*tag- a
XX /product- "TNRL3"
XX MO9911791-A2.
XX 11-MAR-1999.
XX 04-SEP-1998; 98WO-US18393.
```

```
XX 05-SEP-1997; 97US-0924634.
PR (UNIT ) UNIT WASHINGTON.
PA Chaudhary PM;
XX WPI: 1999-205191/17.
XX P-PSDB; AAM93590.
DR New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX Example VII; Fig 13A: 156pp: English.
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX their active fragments. APO4 is useful for diagnosing prostate cancer
XX by determining levels of APO4 in an individual. Prostate cancer can also
XX be treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selective
XX binding agents, useful in diagnosis/treatment of disease by binding of
XX agents to the polypeptide/active fragment which is extracellular, or
XX expressed on the cell surface. The binding is preferably performed in
XX vivo. APO4 polypeptides/active fragments are also useful for screening
XX for agonists and antagonists by binding and observing the change in APO4
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/active
XX fragments and APO4 signal transducer molecules that specifically interact
XX with a cytoplasmic domain of APO4 and detecting a change in level of APO4
XX activity. The method is performed in vivo or in vitro. APO polypeptides
XX are all useful as immunogens for preparing antibodies. APO4 is also
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell line
XX MCF-7, and induced apoptosis.
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;
Query Match 57.1%; Score 784; DB 20; Length 1030;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 229 GTGAGTTGGGAGCGGGGCGATCGTCCGCCAGAGACCTGCCAGAGAGAGCTGGG 288
DB 1 GTGAGTTGGGAGCGGGGCGATCGTCCGCCAGAGACCTGCCAGAGAGAGCTGGG 60
OY 289 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
DB 61 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
OY 349 GCGCTTCTGAGACCGAGTCTGAGGCTGCGAGAAAGTCACTTAAAGCGGAGAAACA 408
DB 121 GCGCTTCTGAGACCGAGTCTGAGGCTGCGAGAAAGTCACTTAAAGCGGAGAAACA 180
OY 409 CCGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
DB 181 CCGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
OY 469 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
DB 241 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 529 AGCCCTGCGGCTACACCGCCAGATCGGGAGTTATAGTACCCGGGCTGGCTTAC 588
DB 301 AGCCCTGCGGCTACTACCGCCAGATCGGGAGTTATAGTACCCGGGCTGGCTTAC 360
OY 589 TACCTGTACTGTACAGTGTACCTTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
DB 361 TACCTGTACTGTACAGTGTACCTTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
OY 649 GTGTGATGTGAGTGTGCTGGCCCTGCGCTGCGAGAGAAATTCAGGCCACTGGCGAGT 708
```

```

Db 421 CTGGTGGATGGTGTGGTGGCCCTCGCTGGAGGAAATTCACGACACGCGCCAGT 480
QY 709 TCCCTCGGGCCCGACAGCTCCGCTTGGCCAGGTGTCTGGGGCTGTGGCCCTGGGCGAGG 768
Db 481 TCCCTCGGGCCCGACAGCTCCGCTTGGCCAGGTGTCTGGGGCTGTGGCCCTGGGCGAGG 540
QY 769 TCCCTCGGGCCCGACAGCTCCGCTTGGCCAGGTGTCTGGGGCTGTGGCCCTGGGCGAGG 828
Db 541 TCCCTCGGGCCCGACAGCTCCGCTTGGCCAGGTGTCTGGGGCTGTGGCCCTGGGCGAGG 600
QY 829 TACTTGGAGCTTCTCCAGCTTCTACATGAGGGCCCTGGCTCTCCACAGTCTCCAGGCT 888
Db 601 TACTTGGAGCTTCTCCAGCTTCTACATGAGGGCCCTGGCTCTCCACAGTCTCCAGGCT 660
QY 889 GCGGCTCCCTCCAGAGCTCTCTGGGACCCGGCTCCCTTGGCCCACTTCCAGGCT 948
Db 661 GCGGCTCCCTCCAGAGCTCTCTGGGACCCGGCTCCCTTGGCCCACTTCCAGGCT 720
QY 949 CTTTGTCTCCAGAGCTCCCTCTCTAGAGGCTGGCTGGCTGTCTTCCAGTGTTCCTCA 1008
Db 721 CTTTGTCTCCAGAGCTCCCTCTCTAGAGGCTGGCTGGCTGTCTTCCAGTGTTCCTCA 780
QY 1009 TCCCACTAATATACAGTATTCCTTATCTTACAACTCCCGACCGCCACT 1063
Db 781 TCCCACTAATATACAGTATTCCTTATCTTACAACTCCCGACCGCCACT 835

```

## RESULT 9

AAS03964 ID AAS03964 standard; DNA: 898 BP.

AAS03964:

26-SEP-2001 (first entry)

Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.

TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 rheobiosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 corneal graft neovascularisation; psoriasis; metastatic condition;  
 malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 preneoplastic condition; myocardial angiogenesis; wound granulation;  
 scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 fusion protein.

Homo sapiens.

OS Synthetic.

Key Location/Qualifiers

52..873  
 /tag- a  
 /product- "fusion protein comprising a growth hormone  
 leader, a leucine zipper multimerisation  
 domain, and human TWEAK extracellular  
 domain"

MO200145730-A2.

28-JUN-2001.

19-DEC-2000: 2000MO-US34755.

20-DEC-1999: 99US-0172878.

PR 10-MAY-2000: 2000US-0203347.

(IMMV ) IMMUNEX CORP.

Wiley SR.

XX WP1: 2001-417975/44.  
 DR P-PSDB: AAU03499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 PS  
 XX Example 1: Page 39-40; 46pp; English.

The sequence represents a DNA from the expression vector  
 pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth  
 hormone leader, a leucine zipper multimerisation domain, and the  
 extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rheobiosis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and  
 CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.

Sequence 898 BP: 187 A; 266 C; 267 G; 178 T; 0 other:

Query Match 45.5%; Score 625; DB 22; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-281;  
 Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 232 AGTTGGGAGACCGGATCGCTCTGCGCCAGAGAGCTGCCAGAGAGCTGTGGCA 291
Db 250 AGTTGGGAGACCGGATCGCTCTGCGCCAGAGAGCTGCCAGAGAGCTGTGGCA 309
QY 292 GAGGAGAGACGAGACCGGCTGCGAATGCCAGACAGAAAGCCAGAGCTCGG 351
Db 310 GAGGAGAGACGAGACCGGCTGCGAATGCCAGACAGAAAGCCAGAGCTCGG 369
QY 352 CTTTCTGAAACGACTAGTTGGCTTCGCAAGAGTCACTAAAGCCGGAAACAGG 411
Db 370 CTTTCTGAAACGACTAGTTGGCTTCGCAAGAGTCACTAAAGCCGGAAACAGG 429
QY 412 GCTCGAAGAGGATCGGACCGCATTTATGAAGTTCATCGACAGCTGGACAGGAGG 471
Db 430 GCTCGAAGAGGATCGGACCGCATTTATGAAGTTCATCGACAGCTGGACAGGAGG 489
QY 472 CAGGACAGGTGTGAGCGGAGACAGTGTGGTGAGGAGAGCCAGATCAACAGCTCAGC 531
Db 490 CAGGACAGGTGTGAGCGGAGACAGTGTGGTGAGGAGAGCCAGATCAACAGCTCAGC 549
QY 532 CCTGTGGCTACAAACCCGACAGATCGGGAGTTTATATCACCCGGGTGGGCTCTACTAC 591
Db 550 CCTGTGGCTACAAACCCGACAGATCGGGAGTTTATATCAACCCGGGTGGGCTCTACTAC 609
QY 592 CTGTACTGTAGGTGACACTTTGATGAGGAGAGGCTGTACCTGAGAGCTGAGCTTCTG 651
Db 610 CTGTACTGTAGGTGACACTTTGATGAGGAGAGGCTGTCTACTGAGAGCTGAGCTTCTG 669
QY 652 GTGAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
Db 670 GTGAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 712 CTGAGGCGCCAGCTCCGCTCTGCGAGGTGTCTGAGGTGTCTGAGGTGTCTGAGGTGTCTG 771

```

DB 730 CTCGGGCCCCAGCTCCGCTCTGCCAGGCTGTGGCTGTGGCCCTCGCCAGGAGTCC 789  
QY 772 TCCCTGGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTCCCTCTCTCACTAC 831  
DB 790 TCCCTGGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTCCCTCTCTCACTAC 849  
QY 832 TTCGACTCTTCCAGTTTCACCTAG 856  
DB 850 TTCGACTCTTCCAGTTTCACCTAG 874  
RESULT 10  
ABK29540  
ID ABK29540 standard; cDNA; 195 BP.  
AC ABK29540;  
DT 23-APR-2002 (first entry)  
XX  
XX  
DE Colon adenocarcinoma-specific cDNA #66.  
XX  
XX Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.  
OS Homo sapiens.  
XX  
XX MO200196389-A2.  
PN  
XX 20-DEC-2001.  
PD  
XX  
XX 07-JUN-2001; 2001WO-US18574.  
PF  
XX 09-JUN-2000; 2000US-210667P.  
PR 22-NOV-2000; 2000US-252614P.  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
PI Meagher MJ, King GE, Xu J, Secret H;  
XX  
XX WPI: 2002-098052/13.  
DR  
XX  
XX  
PT New isolated polynucleotide encoding a polypeptide comprising a portion  
PT of colon tumour protein, for detection, diagnosis and therapy of human  
PT colon cancer -  
XX  
XX  
PS Claim 1; Page 133; 211pp; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) encoding a  
CC polypeptide (II) comprising a portion of a colon tumour protein. A new  
CC oligonucleotide (III) that hybridises to (I) is useful for  
CC determining the presence of a cancer in a patient. (II) or antigen  
CC presenting cells expressing (I) are useful for stimulating and/or  
CC expanding T cells specific for a tumour protein, by contacting T cells  
CC with (I), (II) or antigen-presenting cells that express (I). (I), (II),  
CC or antigen presenting cells that express (II) are useful for treating  
CC colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated  
CC from a patient with (I), (II), or antigen presenting cells that express  
CC (II), so that T cells proliferate, and administering to the patient an  
CC effective amount of the proliferated T cells, thus inhibiting the  
CC development of a cancer in the patient. A new composition is useful for  
CC stimulating an immune response in a patient. (I) or (II) is useful in  
CC vaccines and pharmaceutical compositions for prevention and treatment of  
CC colon cancer and for the diagnosis and monitoring of the cancers. (I),  
CC (II) or an antibody against (II) is useful for detection, diagnosis and/  
CC or therapy of human colon cancer. (I) is useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of ribozyme  
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-  
CC ABK29831 represent human colon adenocarcinoma-specific cDNA sequences of  
CC the invention.  
XX  
XX Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;

Query Match 13.0%; Score 179; DB 24; Length 195;  
Best Local Similarity 100.0%; Pred. No. 2.6e-73;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1146 CATTTGTTCACTGACTCTGTGGGCGACAGATGGGTCCAGAAACCCCACTTCAGGCACT 1205  
DB 17 CATTTGTTCACTGACTCTGTGGGCGACAGATGGGTCCAGAAACCCCACTTCAGGCACT 76  
QY 1206 AAGAGGGGCTGACCTGGGCGGAGCAAGGAGACTGGGCTTAGGCGAGAGTTCC 1265  
DB 77 AAGAGGGGCTGACCTGGGCGGAGCAAGGAGACTGGGCTTAGGCGAGAGTTCC 136  
QY 1266 AATGTGAGGGGCGAGAAACAAGACAGCTCCCTTGAGATTCCTGTGATTTT 1324  
DB 137 AATGTGAGGGGCGAGAAACAAGACAGCTCCCTTGAGATTCCTGTGATTTT 195  
RESULT 11  
AAT22190  
ID AAT22190 standard; cDNA to mRNA; 282 BP.  
AC AAT22190;  
XX  
XX  
XX 27-AUG-1996 (first entry)  
DT  
XX  
XX Human gene signature HUMGS03761.  
DE  
XX  
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9514772-A1.  
PN  
XX  
XX 01-JUN-1995.  
PD  
XX  
XX 11-NOV-1994; 94WO-JP01916.  
PE  
XX  
XX 12-NOV-1993; 93JP-0355504.  
PR  
XX  
XX (MATS/) MATSUBARA K.  
PA (OKUB/) OKUBO K.  
XX  
XX Matsubara K, Okubo K;  
XX  
XX WPI: 1995-206931/27.  
DR  
XX  
XX  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
XX  
XX  
PS Claim 1; Page 1067; 2245pp; Japanese.  
XX  
XX A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in AAT19001-T26837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.  
XX  
XX Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;

Query Match 5.5%; Score 76; DB 16; Length 282;  
Best Local Similarity 100.0%; Pred. No. 2.8e-25;



Matches 76: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 CATTGTCTCACTGCTGTGGGCAAGATGGGTCCAGAACCCACTTCAGGCACT 1205  
|||||  
Db 36 CATTGTCTCACTGCTGTGGGCAAGATGGGTCCAGAACCCACTTCAGGCACT 95  
|||||

OY 1206 AAGAGGGCTGAGCT 1221  
|||||

Db 96 AAGAGGGCTGAGCT 111  
|||||

RESULT 12  
ABN41049  
ID ABN41049 standard; DNA; 60 BP.

XX  
AC ABN41049;  
XX  
DT 15-JUL-2002 (first entry)  
XX

DE Human spliced transcript detection oligonucleotide SEQ ID NO:13797.  
XX  
KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PE 20-JUL-2001; 2001WO-IB01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX

PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

XX  
XX  
PS Example 1; SEQ ID 13797; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
XX messenger RNAs that populate a (sub-)transcriptome, where the  
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
XX transcription units that populate a genome. The library comprises  
XX several oligonucleotides, each capable of hybridizing selectively to a  
XX set of messenger RNAs transcribed from a given transcription unit of  
XX the genome, which encodes one or more messenger RNA splice variants.  
XX The oligonucleotide libraries are useful for detecting mRNAs from a  
XX biological sample, in expression profiling studies, in qualitatively or  
XX quantitatively characterizing the corresponding transcriptome, and in  
XX detecting RNA transcripts and splice variants of human or animal  
XX transcriptomes. The libraries may also be used as specialised mini  
XX libraries to detect transcripts of a sub-transcriptome under a  
XX particular biological or pathological state, and so allowing the  
XX detection of tissue- and pathology-specific genes such as those genes  
XX only expressed in specific tissue under a specific pathological  
XX condition; to detect developmental specific genes; and to detect RNA  
XX transcripts and splice variants of a transcriptome of a patient suffering  
XX from a particular disorder. ABN27253 to ABN5589 represent  
XX oligonucleotide sequences from rats, humans and mice, which are used in  
XX the exemplification of the present invention.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 60 BP; 17 A; 15 C; 15 G; 13 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8 6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1263 CCCAATGTGAGGGCGAAGAACAGACAGCTCCTCCTTGAGAAATCCCTGTGATTT 1322  
|||||  
Db 1 CCCAATGTGAGGGCGAAGAACAGACAGCTCCTCCTTGAGAAATCCCTGTGATTT 60  
|||||

RESULT 13  
ABN58591  
ID ABN58591 standard; DNA; 60 BP.

XX  
AC ABN58591;  
XX  
DT 15-JUL-2002 (first entry)  
XX

DE Human spliced transcript detection oligonucleotide SEQ ID NO:31339.  
XX  
KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PE 20-JUL-2001; 2001WO-IB01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX

PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

XX  
XX  
PS Example 1; SEQ ID 31339; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
XX messenger RNAs that populate a (sub-)transcriptome, where the  
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
XX transcription units that populate a genome. The library comprises  
XX several oligonucleotides, each capable of hybridizing selectively to a  
XX set of messenger RNAs transcribed from a given transcription unit of  
XX the genome, which encodes one or more messenger RNA splice variants.  
XX The oligonucleotide libraries are useful for detecting mRNAs from a  
XX biological sample, in expression profiling studies, in qualitatively or  
XX quantitatively characterizing the corresponding transcriptome, and in  
XX detecting RNA transcripts and splice variants of human or animal  
XX transcriptomes. The libraries may also be used as specialised mini  
XX libraries to detect transcripts of a sub-transcriptome under a  
XX particular biological or pathological state, and so allowing the  
XX detection of tissue- and pathology-specific genes such as those genes  
XX only expressed in specific tissue under a specific pathological  
XX condition; to detect developmental specific genes; and to detect RNA  
XX transcripts and splice variants of a transcriptome of a patient suffering  
XX from a particular disorder. ABN27253 to ABN5589 represent  
XX oligonucleotide sequences from rats, humans and mice, which are used in  
XX the exemplification of the present invention.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 60 BP; 17 A; 15 C; 15 G; 13 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1263 CCCAAATGTGAGGGCGGCAACAGACAGCTCCTCCCTTGAGAAATTCCTGTGGATTT 1322  
DB 1 CCCAAATGTGAGGGCGGCAACAGACAGCTCCTCCCTTGAGAAATTCCTGTGGATTT 60

## RESULT 14

ABN58593  
ID ABN58593 standard; DNA; 60 BP.

XX ABN58593;

XX 15-JUL-2002 (first entry)

XX Human spliced transcript detection oligonucleotide SEQ ID NO:31341.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-1B01903.

XX 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

XX Example 1; SEQ ID 31341; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridizing selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterizing the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialized mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN59589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 60 BP; 12 A; 19 C; 10 G; 19 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 979 GGCTCCTGGGCGCTTCACGTTTCCATCCCAATAAATACAGTATTCACACTCTTA 1038  
DB 1 GGCTCCTGGGCGCTTCACGTTTCCATCCCAATAAATACAGTATTCACACTCTTA 60

## RESULT 15

ABN58848  
ID ABN58848 standard; DNA; 60 BP.

XX ABN58848;

XX 15-JUL-2002 (first entry)

XX Human spliced transcript detection oligonucleotide SEQ ID NO:31596.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-1B01903.

XX 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

XX Example 1; SEQ ID 31596; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridizing selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterizing the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialized mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN59589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 60 BP; 13 A; 16 C; 17 G; 14 T; 0 other;

Query Match

Best Local Similarity 4.4%; Score 60; DB 24; Length 60;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1145 GCATTGTGTTCACTGTCTGTGGCAAGATGGGTCCAGAAGACCCCACTTCAGGCAC 1204

Db 1 GCATTGTGTTCACTGTCTGTGGCAAGATGGGTCCAGAAGACCCCACTTCAGGCAC 60

Search completed: May 8, 2003, 12:20:10  
Job time : 347 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 10:21:50 ; Search time 3521 Seconds

(without alignments)  
11348.524 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373  
Sequence: 1 atgcatgttgtagcttga.....gacaaatgttgtaaatg 1373

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenDbml:  
1: gb\_da:  
2: gb\_hlg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vl:  
15: em\_da:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_om:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vl:  
30: em\_hlg\_hum:  
31: em\_hlg\_inv:  
32: em\_hlg\_other:  
33: em\_hlg\_mus:  
34: em\_hlg\_pln:  
35: em\_hlg\_rtd:  
36: em\_hlg\_mam:  
37: em\_hlg\_vtl:  
38: em\_sy:  
39: em\_hlgo\_hum:  
40: em\_hlgo\_mus:  
41: em\_hlgo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1285	93.6	1306	9 AF030099	AF030099 Homo sapi
2	1172	85.4	1353	6 AX201324	AX201324 Sequence
3	1172	85.4	1368	9 AF055872	AF055872 Homo sapi
4	958	69.8	1236	6 AR140407	AR140407 Sequence
5	796	58.0	1651	6 BC019047	BC019047 Homo sapi
6	625	45.5	898	6 AX180714	AX180714 Sequence
7	620	45.2	177703	2 AC016876	AC016876 Homo sapi
8	435	31.7	215795	2 AC127470	AC127470 Pan trogl
9	179	10.6	195	6 AX379024	AX379024 Sequence
10	145	10.6	177555	2 AC130192	AC130192 Sus scrof
11	119	8.7	138792	2 AC119115	AC119115 Rattus no
12	119	8.7	203083	2 AC069459	AC069459 Mus muscu
13	119	8.7	234182	10 AL603707	AL603707 Mouse DNA
14	59	4.3	184026	2 AC098923	AC098923 Rattus no
15	58	4.2	135533	2 AC126921	AC126921 Bos tauru
16	58	4.2	161428	2 AC126925	AC126925 Canis fam
17	50	3.6	50	6 AX201395	AX201395 Sequence
18	50	3.6	203281	2 AC126237	AC126237 Canis fam
19	48	3.5	164504	2 AC129071	AC129071 Pan trogl
20	46	3.4	1168	10 AF030100	AF030100 Mus muscu
21	34	2.5	139405	2 AC126239	AC126239 Felis cat
22	26	1.9	26	6 AX201459	AX201459 Sequence
23	26	1.9	153	6 AX322295	AX322295 Sequence
24	26	1.9	281	6 AX209050	AX209050 Sequence
25	26	1.9	391	6 AX198506	AX198506 Sequence
26	26	1.9	418	9 BC032480	BC032480 Homo sapi
27	26	1.9	439	6 AX330518	AX330518 Sequence
28	26	1.9	439	6 AX410097	AX410097 Sequence
29	26	1.9	483	6 AX381620	AX381620 Sequence
30	26	1.9	487	6 AX408082	AX408082 Sequence
31	26	1.9	531	6 AX381939	AX381939 Sequence
32	26	1.9	2183	9 HSM80267	HSM80267 Homo sapi
33	26	1.9	110293	9 HSDJ6863	AL049766 Homo sapi
34	26	1.9	159020	2 AC027023	AC027023 Homo sapi
35	26	1.9	203281	2 AC126237	AC126237 Canis fam
36	23	1.7	23	6 AX201458	AX201458 Sequence
37	23	1.7	160340	2 AL845513	AL845513 Danio rer
38	23	1.7	185279	2 AC093996	AC093996 Rattus no
39	22	1.6	1173	1 AF182516	AF182516 Pseudomon
40	22	1.6	2942	10 MM0250723	AJ250723 Mus muscu
41	22	1.6	10989	10 AF4501854	AF450141 Mus muscu
42	22	1.6	41037	2 AC017476	AC017476 Drosophila
43	22	1.6	49536	2 AC130967	AC130967 Rattus no
44	22	1.6	58107	2 AC106343	AC106343 Rattus no
45	22	1.6	62952	2 AC111786	AC111786 Rattus no

# ALIGNMENTS

RESULT 1	AF030099	1306 bp	mrna	linear	PRI 20-DEC-1997
LOCUS	AF030099				
DEFINITION	Homo sapiens TWEAK mRNA, complete cds.				
ACCESSION	AF030099				
VERSION	AF030099.1	GI:2707218			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1306)				
AUTHORS	Chicheportiche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,				
	Hession, C., Garcia, I., and Browning, J.L.				
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family				

that weakly induces apoptosis  
J. Biol. Chem. 272 (51), 32401-32410 (1997)  
09070415









[illegible]

Query Match	45.5%;	Score 625;	DB 6;	Length 898;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 625;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
BASE COUNT	187 a	266 c	267 g	178 t
ORIGIN	/db, xref="GI:15132571" /translation="MATSRKSLLEAFLLPMLQBSATSSDMKOIENKIEPIES KIHLENIATIKLIGERTSSLSRASIQAEPQAEIYAEEDQDSELTNPQTEES ODPAFENLRILVRPRRSAPKGRKTRARRAIAHVEYHPRGDQAQADVTSGMEES RINSSPLRYNNOIGEIFVTRAGLYYLICQVHPDEGKAVYLKDLLVGVIALKCLEE FSVAASSIGPOLRICOVSGLLALRPSSLRIRTLPMHLKAAPFLVFGLFOVH" 266 c 267 g 178 t			
232	ACCTTTGGGGAGCGGGGACATCGCTGTCCGCCGACAGACCTGCCACGAGAGAGCTGTGGCA	291		
250	ACATTGGGGAGCGGGGACATCGCTGTCCGCCGACAGACCTGCCACGAGAGAGCTGTGGCA	309		
292	GAGGAGACACGAGACCCGTCGGAACATGAAATCCCCACAGAGAAGAAAGCAGAGATCCTGCG	351		
310	GAGGAGACACGAGACCCGTCGGAACATGAAATCCCCACAGAGAAGAAAGCAGAGATCCTGCG	369		
352	CCTTTCCCTGAACCCGACTAGTTGGCGCTGGCAGAACTGCACCTAAAGCCGGAAAAACACGG	411		
370	CTTTCTCTGAACCCGACTAGTTGGCGCTGGCAGAACTGCACCTAAAGCCGGAAAAACACGG	429		
412	GCTCGAAGAGCGATGCGACGCCATATGAAGTTCAATCCAGACCTGGACAGAGAGAGCG	471		
430	GCTCGAAGAGCGATGCGACGCCATATGAAGTTCAATCCAGACCTGGACAGAGAGAGCG	489		
472	CAGGAGAGTGTGGACGGGACAGTAGTGGCTGGGAGAAAGCCAGATACACAGCTCCAGC	531		
490	CAGGAGAGTGTGGACGGGACAGTAGTGGCTGGGAGAAAGCCAGATACACAGCTCCAGC	549		
532	CCTCTGCCCTCAACCCGCGAGATCGGGAGATTATAGCACCCGGGCGTGGGCTCTACTAC	591		
550	CCTCTGCCCTCAACCCGCGAGATCGGGAGATTATAGCACCCGGGCGTGGGCTCTACTAC	609		
592	CTGTACTGTCAAGGTGACACTTTGATGAGGGGAAAGGCTGTCTACTGTAAAGCTGGACTGTG	651		
610	CTGTACTGTCAAGGTGACACTTTGATGAGGGGAAAGGCTGTCTACTGTAAAGCTGGACTGTG	669		
652	GTCGATGCTGTGCTGGGCGGCTGCGCTGGAGAGAAATCTCAGCACATCGGGCCAGTTCC	711		
670	GTGATGTGTGTGCTGGGCGGCTGCGCTGGAGAGAAATCTCAGCACATCGGGCCAGTTCC	729		
712	CTCGGGCCCCCAGCTCCGCGCTGTCCAGAGTGTCTGGGGCTGTGGCCCTCGGCCACAGGGTCC	771		
730	CTCGGGCCCCCAGCTCCGCGCTGTCCAGAGTGTCTGGGGCTGTGGCCCTCGGCCACAGGGTCC	789		
772	TCCCTCGGGATCCGACACCTCTCCCTGGGGCCATCTCAAGGCTGCCCTTCTCACTAC	831		
790	TCCCTCGGGATCCGACACCTCTCCCTGGGGCCATCTCAAGGCTGCCCTTCTCACTAC	849		
832	TTCCGACACTTCCAGAGTTCACCTGAG	856		
850	TTCCGACACTTCCAGAGTTCACCTGAG	874		
RESULT 7	AC016876/c 177703 bp DNA linear HTG 06-AUG-2002			
LOCUS	AC016876 Homo sapiens clone RP11-186B7, *** SEQUENCING IN PROGRESS ***, 10			
DEFINITION	unordered pieces.			
ACCESSION	AC016876			
VERSION	AC016876.5 GI:21313830			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULINOP; HTGS_ACTIVEFIN.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 177703)			
AUTHORS	Bliren,B., Nusbaum,C. and Lander,E.			
TITLE	Homo sapiens chromosome, clone RP11-186B7			

**JOURNAL  
REFERENCE  
AUTHORS**

Unpublished  
2 (bases 1 to 177703)  
Barren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bonkhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gaidyna, S., Grant, G., Hagos, B., Heiford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Romann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.-J., Zimmer, A., and Zody, M.

TITLE	JOURNAL	REFERENCE	AUTHORS
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
3. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
4. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
5. The Effect of Surface Area on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.

Direct Submission  
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 120 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 177703)  
Bairren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, V., Collymore, A.,  
Cook, A., Cooke, P., Detrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galaan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norby, C., Norman, C. H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Risse, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schuppach, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, K., Wyman, D., Young, G., Zainoun, J.,  
Ziemek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

**Direct Submission**  
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 2, 2002 this sequence version replaced gi:154212189.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>







(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 1022: contig of 1022 bp in length
* 1023 1122: gap of unknown length
* 1123 2394: contig of 1272 bp in length
* 2395 2494: gap of unknown length
* 2495 3663: contig of 1169 bp in length
* 3664 3763: gap of unknown length
* 3764 4879: contig of 1116 bp in length
* 4880 4979: gap of unknown length
* 4980 6787: contig of 1808 bp in length
* 6788 8192: gap of unknown length
* 8193 8292: gap of unknown length
* 8293 9461: contig of 1169 bp in length
* 9462 9561: gap of unknown length
* 9562 11435: contig of 1874 bp in length
* 11436 11535: gap of unknown length
* 11536 13535: contig of 2000 bp in length
* 13536 13635: gap of unknown length
* 13636 17323: contig of 3688 bp in length
* 17324 17424: gap of unknown length
* 17424 20406: contig of 2983 bp in length
* 20407 20507: gap of unknown length
* 20507 23445: contig of 2939 bp in length
* 23446 23545: gap of unknown length
* 23546 25600: contig of 2055 bp in length
* 25601 25701: gap of unknown length
* 25701 28049: contig of 2349 bp in length
* 28050 28149: gap of unknown length
* 28150 30474: contig of 2325 bp in length
* 30475 30574: gap of unknown length
* 30575 35498: contig of 4924 bp in length
* 35499 35598: gap of unknown length
* 35599 39284: contig of 3686 bp in length
* 39285 39384: gap of unknown length
* 39385 41970: contig of 2586 bp in length
* 41971 42070: gap of unknown length
* 42071 45659: contig of 3589 bp in length
* 45660 45759: gap of unknown length
* 45760 50332: contig of 4573 bp in length
* 50333 50432: gap of unknown length
* 50433 54695: contig of 4263 bp in length
* 54696 54795: gap of unknown length
* 54796 58289: contig of 3494 bp in length
* 58290 58389: gap of unknown length
* 58390 63599: contig of 5210 bp in length
* 63600 63699: gap of unknown length
* 63700 70305: contig of 6606 bp in length
* 70306 70405: gap of unknown length
* 70406 76123: contig of 5718 bp in length
* 76124 76223: gap of unknown length
* 76224 84961: contig of 8738 bp in length
* 84962 85061: gap of unknown length
* 85062 93614: contig of 8553 bp in length
* 93615 93714: gap of unknown length
* 93715 103352: contig of 9638 bp in length
* 103353 103452: gap of unknown length
* 103453 110299: contig of 6847 bp in length
* 110300 110399: gap of unknown length
* 110400 116180: contig of 5781 bp in length
* 116181 116280: gap of unknown length
* 116281 125484: contig of 9204 bp in length
* 125485 125584: gap of unknown length
* 125585 138792: contig of 13208 bp in length.
```

## FEATURES

Location/Qualifiers

## source

1. 138792

/organism="Rattus norvegicus"

/db\_xref="taxon:10116"

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others

## ORIGIN

Query Match 8.7%; Score 119; DB 2; Length 138792;

Best Local Similarity 100.0%; Pred. No. 1,4e-34;

Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 51 ATCCCTCGGATCCGAGATGGGGGCGGTGAGGAGGACAGCCCGCCCGCATGGC 110

Db 129318 ATCCCTCGGATCCGAGATGGGGGCGGTGAGGAGGACAGCCCGCCCGCATGGC 129377

Oy 111 CGCCGTCGAGCCAGAGCGAGGCGCGCGCGGGGAGCCCGGACCGCCCTGTCG 169

Db 129378 CGCCGTCGAGCCAGAGCGAGGCGCGCGCGGGGAGCCCGGACCGCCCTGTCG 129436

## RESULT 12

AC069459/c

LOCUS

DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT

ACCESSION AC069459.23 GI:14547768

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 203083) Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

Metzger, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,

Dederich, D., Thomas, S., Okunou, G., Carllock, C., Garner, T.,

Addison, S., Pace, A., Williams, G., Bonaldi, D., Brooks, A., Brown, J.,

Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,

Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,

Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,

Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M.,

Hosak, H., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,

Kovar, C., Liu, J., Liu, W., Louised, H., Lozano, R., Martin, R.,

Maesey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,

Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogul, M., Parish, B.,

Perez, L., Reller, D., Say, J., Shen, H., Vasquez, L., Wallington, S.,

Williamson, A., Wrenford, G., Zhou, X., Bouck, J., Hodgson, A.,

Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,

Worley, K. and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 203083)

Worley, K.C.

Direct Submission

Submitted (31-MAY-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jun 25, 2001 this sequence version replaced gi:12621364.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: MAF0

Center clone name: RP23-168P5

----- Summary Statistics

Sequencing vector: M13; 108821

Chemistry: Dye-terminator Big Dye 52% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 212648 bases at least Q40

Consensus quality: 218902 bases at least Q30

Consensus quality: 222384 bases at least Q20



Estimated insert size: 21056; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.bjsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.bjsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 62152: contig of 62152 bp in length  
\* 62153 62252: gap of unknown length  
\* 62253 118772: contig of 56520 bp in length  
\* 118773 118872: gap of unknown length  
\* 118873 148924: contig of 30052 bp in length  
\* 148925 149024: gap of unknown length  
\* 149025 167331: contig of 18207 bp in length  
\* 167332 189907: gap of unknown length  
\* 189908 190007: contig of 22576 bp in length  
\* 190008 196537: gap of unknown length  
\* 196538 203083: contig of 6446 bp in length.  
\* 196538 203083: contig of 6446 bp in length.

FEATURES  
source  
1. 203083  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others

ORIGIN

Query Match 8.7%; Score 119; DB 2; Length 203083;  
Best Local Similarity 100.0%; Pred. No. 1.4e-54;  
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ATCCCTGGGTCCCGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 110  
|||||  
Db 51733 ATCCCTGGGTCCCGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 51674

QY 111 CGCCCGTGGAGACAGAGGCGGAGGCGCCCGGGGAGACCGCGCCCTGCTGG 169  
|||||  
Db 51673 CGCCCGTGGAGACAGAGGCGGAGGCGCCCGGGGAGACCGCGCCCTGCTGG 51615

RESULT 13  
AL603707/c 234182 bp DNA linear ROD 17-NOV-2001  
LOCUS  
DEFINITION Mouse DNA sequence from clone RP23-422L16 on chromosome 11,  
complete sequence.  
ACCESSION AL603707  
VERSION AL603707.5 GI:17017790  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 234182)

REFERENCE 1 (bases 1 to 234182)  
AUTHORS Pearce, A.  
TITLE Direct Substitution  
JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humay@anger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Nov 20, 2001 this sequence version replaced gi:16605765.

COMMENT During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: EM: EMBL; SW:  
SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information on the WORMPEP  
database can be found at  
<http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-422L16 is  
from the RPECI-23 Mouse PAC library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VCFOR: PBA063.6  
This sequence is the entire insert of clone RP23-422L16.

FEATURES  
source  
1. 234182  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-422L16"  
/clone\_lib="RPECI-23"  
/complement="84050..84131"  
/note="Sequence from uni-directional primer reads and dGTP  
big dye terminator reads only."

BASE COUNT 59310 a 56824 c 57519 g 60529 t

ORIGIN

Query Match 8.7%; Score 119; DB 10; Length 234182;  
Best Local Similarity 100.0%; Pred. No. 1.4e-54;  
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ATCCCTGGGTCCCGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 110  
|||||  
Db 75038 ATCCCTGGGTCCCGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 74979

QY 111 CGCCCGTGGAGACAGAGGCGGAGGCGCCCGGGGAGACCGCGCCCTGCTGG 169  
|||||  
Db 74978 CGCCCGTGGAGACAGAGGCGGAGGCGCCCGGGGAGACCGCGCCCTGCTGG 74920

RESULT 14  
AC098923/c 184026 bp DNA linear HTG 12-JUL-2002  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-154B15, \*\*\* SEQUENCING IN PROGRESS  
ACCESSION AC098923  
VERSION AC098923.5 GI:2129926  
KEYWORDS HTGS PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 184026)

REFERENCE 1 (bases 1 to 184026)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., All-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaral, H.C., Aré, J.R., Ayala, M., Banks, T.,  
Barbieri, J., Benton, J., Bimge, K., Blankenburg, K., Bonini, D.,  
Bouck, J., Bowler, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, R.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dedrich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Doutwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, F., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliver,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S., Ogun,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Qules,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Unpublished  
Direct Submission  
2 (bases 1 to 184026)  
Worley,K.C.

Direct Submission  
Submitted (06-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 184026)  
Worley,K.C.

Direct Submission  
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:17973852.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: G10K  
Center clone name: CH230-154B15  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 109715 bases at least Q40  
Consensus quality: 115369 bases at least Q30  
Consensus quality: 121073 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 71 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1385: contig of 1385 bp in length  
\* 1386 1485: gap of unknown length  
\* 1486 2496: contig of 1011 bp in length  
\* 2497 2596: gap of unknown length  
\* 2597 3886: contig of 1290 bp in length  
\* 3887 3986: gap of unknown length  
\* 3987 5552: contig of 1566 bp in length  
\*  
\* 5553 5652: gap of unknown length  
\* 5653 7125: contig of 1473 bp in length  
\* 7126 7225: gap of unknown length  
\* 7226 8566: contig of 1341 bp in length  
\* 8567 8666: gap of unknown length  
\* 8667 9765: contig of 1098 bp in length  
\* 9765 9864: gap of unknown length  
\* 9864 11285: contig of 1421 bp in length  
\* 11286 11385: gap of unknown length  
\* 11386 12407: contig of 1022 bp in length  
\* 12408 12507: gap of unknown length  
\* 12508 13827: contig of 1320 bp in length  
\* 13828 13927: gap of unknown length  
\* 13928 15227: contig of 1300 bp in length  
\* 15228 15327: gap of unknown length  
\* 15328 16441: contig of 1114 bp in length  
\* 16442 16541: gap of unknown length  
\* 16542 17797: contig of 1256 bp in length  
\* 17798 17897: gap of unknown length  
\* 17898 19717: contig of 1820 bp in length  
\* 19718 19817: gap of unknown length  
\* 19818 21264: contig of 1447 bp in length  
\* 21265 21364: gap of unknown length  
\* 21365 22396: contig of 1032 bp in length  
\* 22397 22496: gap of unknown length  
\* 22497 23618: contig of 1122 bp in length  
\* 23619 23718: gap of unknown length  
\* 23719 25080: contig of 1362 bp in length  
\* 25081 25180: gap of unknown length  
\* 25181 26592: contig of 1412 bp in length  
\* 26593 26692: gap of unknown length  
\* 26693 27883: contig of 1191 bp in length  
\* 27884 27983: gap of unknown length  
\* 27984 29261: contig of 1278 bp in length  
\* 29262 29361: gap of unknown length  
\* 29362 31112: contig of 1751 bp in length  
\* 31113 31212: gap of unknown length  
\* 31213 32801: contig of 1589 bp in length  
\* 32802 32901: gap of unknown length  
\* 32902 34318: contig of 1417 bp in length  
\* 34319 34418: gap of unknown length  
\* 34419 36405: contig of 1987 bp in length  
\* 36406 36505: gap of unknown length  
\* 36506 37644: contig of 1139 bp in length  
\* 37645 37744: gap of unknown length  
\* 37744 39164: gap of unknown length  
\* 39165 39264: contig of 1420 bp in length  
\* 39265 41252: gap of unknown length  
\* 41253 41352: contig of 1988 bp in length  
\* 41353 42492: contig of 1140 bp in length  
\* 42493 42592: gap of unknown length  
\* 42593 44098: contig of 1506 bp in length  
\* 44099 44198: gap of unknown length  
\* 44199 46386: contig of 2188 bp in length  
\* 46387 46486: gap of unknown length  
\* 46487 48067: contig of 1581 bp in length  
\* 48068 48167: gap of unknown length  
\* 48168 49678: contig of 1511 bp in length  
\* 49679 49778: gap of unknown length  
\* 49779 51600: contig of 1822 bp in length  
\* 51601 51700: gap of unknown length  
\* 51701 53618: contig of 1918 bp in length  
\* 53619 53718: gap of unknown length  
\* 53719 56000: contig of 2282 bp in length  
\* 56001 56100: gap of unknown length  
\* 56101 58397: contig of 2297 bp in length  
\* 58398 58497: gap of unknown length  
\* 58498 60619: contig of 2121 bp in length  
\* 60619 60718: gap of unknown length  
\* 60719 62515: contig of 1797 bp in length  
\* 62516 65432: gap of unknown length  
\* 65433 65532: contig of 2817 bp in length  
\* 65532: gap of unknown length



```
misc_feature      /note="assembly_fragment"  
90960. .111428  
misc_feature      /note="assembly_fragment"  
111529. .115353  
clone_end:SP6  
vector_side:right"  
BASE COUNT      39178 a 37393 c 36616 g 39259 t 1107 others  
ORIGIN  
Query Match      4.2%; Score 58; DB 2; Length 153553;  
Best Local Similarity 100.0%; Pred. No. 2.4e-20;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 601 CAGTGCACCTTGTGATGAGGGGAAGGCTGTCTACCTGAACCTGGACTTCTGCTGGATG 658  
|||||  
DB 27189 CAGTGCACCTTGTGATGAGGGGAAGGCTGTCTACCTGAACCTGGACTTCTGCTGGATG 27246
```

Search completed: May 8, 2003, 13:40:41  
Job time : 4822 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:27:27 ; Search time 1056.48 Seconds

(without alignments)  
3449.164 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSGLALACGLLVVYSL.....PMHLKAPFLTYGLFQVH 225

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n model -DEV=xlp  
-O=/ggn2.1/USFto\_spool/US09245198/runat\_06052003\_170239\_9628/app\_query.fasta\_1.846  
-DB-EST-QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=100 -THR\_MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09245198 @CGN 1.1.2375 @runat\_06052003\_170239\_9628 -NCPU=6 -ICPU=3  
-NO\_XIPYX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estha:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlin:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estlin:\*  
16: em\_estlin:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_dln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1039	89.4	918	12	BF577781
2	968.5	83.3	731	13	BF577781
3	941.5	81.0	728	13	BF577781
4	882	75.9	777	13	BF577781
5	861	74.1	828	13	BF577781
6	851	74.1	834	13	BF577781
7	857	73.8	861	13	BF577781
8	831.5	71.6	948	14	BF0707185
9	818	70.4	940	14	BQ084231
10	764	65.7	963	14	BQ671259
11	754	64.9	1033	11	AK020909
12	721	62.0	1071	14	BM921213
13	703	60.5	951	14	BQ674188
14	699	60.2	584	10	AM917574
15	670	57.7	445	9	AA221610
16	632.5	54.4	471	9	AA221610
17	629	54.1	650	12	BG404836
18	621	53.4	360	10	BE654876
19	610	52.5	785	13	BF762908
20	561.5	48.3	542	12	BF041509
21	549.5	47.3	404	12	BF044430
22	541	46.6	315	12	BF466521
23	513	44.1	474	13	BF1965174
24	506.5	43.6	493	10	BE307031
25	501.5	43.2	894	13	BF1908274
26	486	41.8	558	13	BM484863
27	477	41.0	581	13	BF1738634
28	473	40.7	538	12	BF821434
29	458	39.4	298	14	BM688946
30	429	36.9	278	10	AM417023
31	428	36.8	342	9	AA637970
32	404.5	34.8	571	12	BF073881
33	404	34.8	345	14	R55379
34	392	33.7	587	12	BG686319
35	373	32.1	400	9	AI152313
36	373	32.1	441	13	BF1967060
37	370	31.8	487	10	AM320117
38	369.5	31.8	698	13	BF1906850
39	354	30.5	260	10	BB596688
40	338	29.1	533	10	BB628951
41	336	28.9	910	12	BG110063
42	335	28.8	456	13	BF1966255
43	316	27.2	687	14	BQ208433
44	287	24.7	531	13	BF182443
45	243.5	21.0	639	10	BB642326

## ALIGNMENTS

RESULT 1  
LOCUS BF577781  
DEFINITION 602092080F1 NCI\_CGAP\_C024 Mus musculus cDNA clone IMAGE:4206595 5',  
mRNA sequence.  
ACCESSION BF577781  
VERSION BF577781.1 GI:11651493  
KEYWORDS  
SOURCE  
ORGANISM  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 918)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)





Qy	65	ValArgProArgArgSerSerAlaProLysGlyIArgLysAlaArgProArgAlaAlaIleAla	84
Db	182	GTGGGCTCCACAAAGGACCTAAAGCCGGAAACACAGAGGCTCGAAGACGATCCCA	241
Qy	85	AlaHisTyrGluValAlaHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly	104
Db	242	GCCCATTTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGACGAGTGTTGGACGG	301
Qy	105	ThrValSerGlyTyrPglGluGluThrLysIleAsnSerSerSerProLeuArgTyrAspArg	124
Db	302	ACAGGAGTGGCTGGGAGGAAGCAGAACTCAACAGCTCCACCCCTCTGGCTACAAACCGC	361
Qy	125	GlnIleGlyGluPheThrValIleLeuArgIleGlyLeuTyrTyrLeuTyrCysGlnValHis	144
Db	362	CAGATCGGGGAGTTATATAGTACACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTGCAC	421
Qy	145	PheAspGluGlyLysAlaValTyrIleuLysLeuAspLeuValAsnGlyValLeuAla	164
Db	422	TTTGATGAGGGGAAAGGCTGTCTACCTGAAGCTGGACCTGTGTGTGATGATGTGTGTGGCC	481
Qy	165	LeuArgCysLeuGluGluPheSerLeuAlaHisAlaAspSerProGlyProGlnLeuArg	184
Db	482	CTGGCTCTCCGGAGGAATCTCACCCCTGGGCAATCTCCCTCGGGCCCCAGCTCCGC	541
Qy	185	LeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerSerLeuArgIleArgThr	204
Db	542	CTCTGCCACGAGTGTCTGGGCTGTGGCCCTGGGGCCAGGTCC-TCCCTCGGATCCGCACAC	600
Qy	205	LeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal	224
Db	601	CTCCCTCGGGCCCATCTCAAGGCTGCCCTTCCTCCACCTCTCGGACCTCTCCAGGTT	660
Qy	225	His 225	
Db	661		
		CAC 663	
RESULT 3			
LOCUS	B1870393		
DEFINITION	B1870393	728 bp	mRNA
	603395641P1 NIH_MGC_90 Homo sapiens	cdna	linear
	mRNA sequence.		EST 11-OCT-2001
VERSION	B1870393		
ACCESSION	B1870393.1	GI:16044066	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 728)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgapds@email.nih.gov">cgapds@email.nih.gov</a>		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: LLM12034 row: a column: 12		
	High quality sequence stop: 728.		
FEATURES			
source	location/Qualifiers		
	1..728		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5405459"		
	/clone_id="NIH_MGC_90"		
	/tissue_type="adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI;		

Site\_2: SalI: cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by life Technologies  
Note: this is a NIH\_MGC library."

Alignment Scores:	
Pred. No.:	2,99e-89
Score:	7287
Percent Similarity:	941.50
Best Local Similarity:	91.86%
Best Local Similarity:	87.7788
Query Match:	81.02%
DB:	13
Length:	7287
Matches:	194
Conservative:	9
Mismatches:	17
Indels:	3
Gaps:	1

US-09-245-198A-2 (1-225) x BI870393 (1-728)

[illegible]



OY 46 GluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeuGlnGluLeuVal 65  
 DB 44 GAACAGATCCCCACACAGAGAAAGCCAGATCTCGCCCTTCTCTGACCGACTAGTT 103  
 OY 66 ArgProArgArgSerAlaProGlyAlaArgValAlaArgProArgArgAlaAlaAla 85  
 DB 104 CGGCTCCAGAGAGTGCACCTTAAGCCGGAACACAGCGCTCGAAGACGATCCAGCC 163  
 OY 86 HisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThr 105  
 DB 164 CATATGAGATTCATCCACGACCTGACAGAGAGCGCAGCGCTGTGACCGGACA 223  
 OY 106 ValSerGlyTyrPglGluGlnThrLysIleAsnSerSerSerProLeuArgTyrAspArgin 125  
 DB 224 TTGATGGCTGGAGAGAACCCAGATCAACACCTCCAGCCCTCGCTACACCGCCAG 283  
 OY 126 IleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 145  
 DB 284 ATCGGGAGATTATAGTACCCCGGCTGCGCTCTACTACTGCTACTGCTGAGCTTT 343  
 OY 146 AspGluGlyValAlaValTyrIleuLysLeuAspLeuValAsnGlyValLeuAlaLeu 165  
 DB 344 GATGAGGGAGAGCTGTCTACTCTGAGAGCTGCTGCTGATGCTGCTGCTGCTG 403  
 OY 166 ArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnArgLeu 185  
 DB 404 CGCTCCCTGGAGAGATTTCTACACCTGCGGCGGCTGCTGCGGCGGCTGCGGCTG 463  
 OY 186 CysGlnValSerGlyLeuLeuProLeuArgProGlySerSerSerLeuArgIleArgThrLeu 205  
 DB 464 TGCCAGGTGTCTGGCTGTGGCCCTGCGGCGGCTGCTGCTGCGGCTGCGGCTG 523  
 OY 206 ProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
 DB 524 CCTGGGCGCATCTCAAGCTGCCCCCTTCTCTACCTACTGAGCTTCTTCAGGTTTCA 583  
 RESULT 6  
 B1766766 834 bp mRNA linear EST 25-SEP-2001  
 LOCUS 603056866F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5206217 5'  
 DEFINITION mRNA sequence.  
 ACCESSION B1766766  
 VERSION B1766766.1 GI:15758344  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 834)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L1AM1517 row: c column: 18  
 High quality sequence stop: 772.  
 Location/Qualifiers  
 1..834  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5206217"  
 /clone\_id="NIH\_MGC\_122"  
 /lab\_host="DH10B"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;  
 Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female  
 spleen, and 20-22 week male spleens. Library is oligo-dT  
 primed and directionally cloned (ecory site is destroyed  
 upon cloning). Average insert size 1.4 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 026. Note:  
 this is a NIH-MGC library."

BASE COUNT 154 a 287 c 223 g 170 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,19e-80 Length: 834  
 Score: 861.00 Matches: 172  
 Percent Similarity: 92.27% Conservative: 7  
 Best Local Similarity: 88.66% Mismatches: 14  
 Query Match: 74.10% Indels: 2  
 DB: 13 Gaps: 0

US-09-245-198A-2 (1-225) x B1766766 (1-834)

OY 33 GlnGluGluLeuThrAlaGluAspArgGluProProGluLeuAsnProGlnThrGlu 52  
 DB 3 CAGAGAGAGCTGGTGGCAGAGAGACAGACAGCCGTCGGAATCAATCCACAGAGA 62  
 OY 53 GlnSerGlnAspValValProPheLeuGluGlnLeuValArgProArgArgSerAlaPro 72  
 DB 63 GAAGCCAGAGATCTCGCCCTTCTCTGAGACAGTGTGGCTGCGGAGAGTGCACCT 122  
 OY 73 LysGlyArgLysAlaArgProArgArgAlaIleAlaAlaHisTyrGlnValHisProArg 92  
 DB 123 AAAGCCGGAAGAAACAGGGCTGGAAGAGGATCGAGCCCATTAAGATTATCCACGA 182  
 OY 93 Pro-GlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluThr 112  
 DB 183 CGGTGACAGCAGCAGGAGCGCAGCAGCTGTGACGCGGACAGTGTGGCTGGAGAGAGC 242  
 OY 112 LysIleAsnSerSerSerProLeuArgTyrAspArginIleGlyGluPheThrVal11 132  
 DB 243 CAGATTCACACCTCCAGCCCTCTGCTACACCGCCAGATGGGAGATTATAGTCA 302  
 OY 132 eArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyr 152  
 DB 303 CCGGCTGGGCTCTACTACTGCTGAGTGCAGTGTGATGAGGGAGAGCTGTCTCA 362  
 OY 152 IleuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGluPhe 172  
 DB 363 CTGAGAGCTGAGCTGTGTGTGATGTGTGCTGCGCCCTGCTGCGAGAAATCTTC 422  
 OY 172 rAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnValSerGlyLeu 192  
 DB 423 AGCCATCGGCGCATTTCCCTGGGCGCCAGCTCCGCTCTGCAAGTGTCTGGCTGTT 482  
 OY 192 uProLeuArgProGlySerSerSerLeuArgIleArgThrLeuProTyrAlaHisLeuLysAl 212  
 DB 483 GGCCCTGGGCGCAGGTTC-TCCCTGGATTCGCAACCTCCCTGGGCGCATCTCAAGC 541  
 OY 212 aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
 DB 542 TGCCCTCTCTCCACCTACTTGAGCTCTTCAGGTTTCA 581  
 RESULT 7  
 AM763237 561 bp mRNA linear EST 04-MAY-2000  
 LOCUS ur70d09.y1 NCI-CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'  
 DEFINITION similar to TR:054907 054907 TMF-RELATED WEAK INDUCER OF APOPTOSIS  
 ;, mRNA sequence.  
 ACCESSION AM763237  
 VERSION AM763237.1 GI:7695174  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
1 (bases 1 to 561)  
AUTHORS  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Other\_ESTS: ur70d09.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[image.llnl.gov/image/html/resources.shtml](http://image.llnl.gov/image/html/resources.shtml)

FEATURES  
source  
MGI:1058389  
Seq primer: -40RP from G1bco  
High quality sequence stop: 433.  
Location/Qualifiers  
1..561  
/organism="Mus musculus"  
/strain="129,C57BL/6J,FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3155633"  
/clone\_id="NCI-CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Salt;  
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by: Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1,77e-80 Length: 561  
Score: 857.00 Matches: 167  
Percent Similarity: 99.40% Conservative: 0  
Best Local Similarity: 99.40% Mismatches: 1  
Query Match: 73.75% Indels: 0  
DB: 10 Gaps: 0

US-09-245-198a-2 (1-225) x AM763237 (1-561)

QY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeu 20  
|||||  
Db 58 GTGCTGAGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 117  
QY 21 GlySerTrpAlaThrLeuSerAlaGlnGlnProSerGlnGlnGlnLeuThrAlaGln 40  
|||||  
Db 118 GGGAGCTGGGCAACGGCTGTGCCAGAGACCTTCTAGAGAGCGTGAAGCAGAGAC 177  
QY 41 ArgArgGluProProGluLeuAsnProGlnThrGlnGlnSerGlnAsnValValProPhe 60  
|||||  
Db 178 CGCCGGAGAGCCCTGAGCTGATCCACACAGAGAAAGCCAGATGTGGTACTTTC 237  
QY 61 LeuGlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArg 80  
|||||  
Db 238 TTGGACACACTAGTCCGGCTCGAAGAGTGTCTTAAGGCGGAGAGGCGGCGCTCGC 297  
QY 81 ArgAlaAlaLeuAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100  
|||||  
Db 298 CGAGCAATTCACGCCCTTGTAGAGTTTCATCTCGGCGACAGAGATGGACACAGCA 357  
QY 101 G1ValAspGlyThrValSerGlyTyrGlnGlnThrLysIleAsnSerSerProLeu 120  
|||||  
Db 358 GGTGTGATGGACAGTGAAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG 417

QY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140  
|||||  
Db 418 CGGTACAGCCGACAGATTTGGGAATATACAGTCATGAGGCTGGGCTCTACTACTGTAC 477  
QY 141 CysGlnValHisPheAspGlyGlyValValTyrLeuLysLeuAspLeuValAsn 160  
|||||  
Db 478 TGTCAAGTGCACCTTGTATGAGGGAAGGCTGTACTTCAAGCTGACCTGTGGTGAAC 537  
QY 161 G1ValLeuAlaLeuArgCysLeu 168  
|||||  
Db 538 GGTGTGCTGGCCCTGCGCTGCTG 561

RESULT 8  
LOCUS BQ707185 948 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens CDNA clone IMAGE:6278608  
5', mRNA sequence.  
ACCESSION BQ707185  
VERSION BQ707185.1 GI:21846084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 948)  
AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Dr. Mark Watson  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LICM2466 row: n column: 17  
High quality sequence start: 24  
High quality sequence stop: 550.  
Location/Qualifiers

FEATURES  
source  
1..948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6278608"  
/clone\_id="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pOTB7; Site:1: XhoI; Site:2:  
EcoRI; CDNA made by oligo-dT priming. directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using zap-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: This is a  
NIH\_MGC library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.9e-77 Length: 948  
Score: 831.50 Matches: 164  
Percent Similarity: 92.31% Conservative: 4  
Best Local Similarity: 90.11% Mismatches: 13  
Query Match: 71.56% Indels: 2  
DB: 14 Gaps: 1

US-09-245-198a-2 (1-225) x BQ707185 (1-948)

QY 44 ProProGluLeuAsnProGlnThrGlnGlnSerGlnAspValValProPheLeuGln 63  
|||||  
Db 3 CCGTGGACACTGATCCCGACAGAGAAAGCAGATCTCGCTTCTTCAACCGA 62  
QY 64 LeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIle 83

```

|||||
Db 63 CTAGTTCGGCCCTCGAAGAGTCACCTAAAGCCGGGAAACACAGGGGCTCGAAGAGCAGATC 122
Oy 84 AAlaAlaHs1TyrgLyuAlaHs1ProArpProGlyGlnAspGlyAlaGlnAlaGlyValAsp 103
Db 123 GCAGGCCCATTTAGAGATTTCATCCAGACCTGACAGAGAGAGAGCAGAGCAGAGCAGAGCAG 178
Oy 104 G1YThrValSerGlyTyrGlnGluThrLysIleAsnSerSerProLeuArgTyrAsp 123
Db 179 GGGACAGTGAAGTGGCTGGAGGAGACCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAC 238
Oy 124 ArgGlnIleGlyLuphTherValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 143
Db 229 CGCCAGATCGGGAGCTTTTACTACACCGGGCTGGGCTCTACTACTGACTGACTGAGGTC 298
Oy 144 HisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValLeu 163
Db 299 CACTTTGATGAGGAGGAGAGGCTCTACCTGAGAGCTGAGCTGCTGATGATGATGATGATG 358
Oy 164 AlaLeuArgCysLeuGlnGluLuphSerAlaThrAlaAlaSerSerProGlyProGlnLeu 183
Db 359 GCCCTGCGCTGCTGAGAGAAATTCAGACCTGCGGCGAGTTCCTCGGCGCCAGCTC 418
Oy 184 ArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArg 203
Db 419 CGCCTCGCCAGAGTGTCTGGGCTGTGGCCCTGGCGCCAGAGCTCTCTCCCTGGGATCCGC 478
Oy 204 ThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln 223
Db 479 ACCCTCCCTGGCCCATCTCAAGAGCTGCGCCCTCTCCTACTGCTGAGACTCTTCAG 538
Oy 224 ValHis 225
Db 539 GTTCAC 544

RESULT 9
LOCUS B0884231 940 bp mRNA linear EST 16-AUG-2002
DEFINITION AGNCOCRT_8682031 lupskl_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6197488 5', mRNA sequence.
ACCESSION B0884231
VERSION B0884231.1 GI:22276239
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13607 row: j column: 17
High quality sequence stop: 453.
Location/Qualifiers
1..940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6197488"
/clone_lib="Lupski_sciatic_nerve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); site_1:"

```

Noti: Site 2: Salt: cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCAGGCTCCG-3' and  
 5'-GACTAGTTCAGATCGCAGCGGCCCTT(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies.

```

BASE COUNT      170 a      325 c      255 g      190 t
ORIGIN
Alignment Scores:
Pred. No.:      5,03e-76      Length:      940
Score:          818.00      Matches:      166
Percent Similarity: 92.02%      Conservative: 7
Best Local Similarity: 88.30%      Mismatches: 11
Query Match:      70.40%      Indels:      4
DB:              14      Gaps:      0

```

```

US-09-245-198a-2 (1-225) x B0884231 (1-940)
Oy 42 ArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAspValAlaProPheLeu 61
Db 1 CAGGACCCGCTGGAGACTGATGCCAGACAGAAAGCAGAGATCTCGGCTTCTC 60
Oy 62 G1uGlnLeuValArgProArgArgSerAlaProGlyGlyArgLysAlaArgProArg 81
Db 61 AACCCACTAGTTCGGCTCGCAGAGAGTGACCTAAAGCCGGGAAACAGCGGCTGAGAGA 120
Oy 82 AlaIleAlaAlaHs1TyrgLyuAlaHs1ProArpProGlyGlnAspGlyAlaGlnAlaGly 101
Db 121 GCGATCGCGCCCATTTATGAACTTCATCCACACCTGGACAGAGAGCAGCCAGCT 180
Oy 102 ValAspGlyThrValSerGlyTyrPglGluThrLysIleAsnSerSerProLeuArg 121
Db 181 GTGAGAGGAGACAGTGAAGTGTGGAGAGAAACAGAAATCAACAGCTCCAGCCCTCGCGC 240
Oy 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
Db 241 TCAACCCGCGAGATCGGGAGGATTTATAGTCACCCGGGCTGCTACTACCTGACTGCT 300
Oy 142 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 301 CAGGTGCACTTTGATGAGAGGAGGAAGCTGTCTACCTGAAGCTGACTGCTGATGATGCT 360
Oy 162 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
Db 361 GTGCTGGCCCTCGCTGCTGGAGGAATCTCAGCCAGCTGCGGAGTCCCTCGGCGCC 420
Oy 182 GlnLeuArgLeuCys-GlnValSerGlyLeuLeuProLeuArgProGlySer-SerLeu 201
Db 421 CAGCTCCGCTCTGCCAGAGTGTGGGCTGTGGCCCTCGCGGCAAGGGGCTCCCTGTC 480
Oy 201 rG1lArGThrLeuProTyrAlaHisLeuLysAlaAla-ProPheLeuThrTyr-PheG 220
Db 481 GATTCGCGACCTCCCTCCGAGCCATCTCAAGAGCTGCCCCCTTCTCACTACTTTGG 540
Oy 220 YLeuPheGlnValHis 225
Db 541 ACTCTTCAGGTTTCAC 556

RESULT 10
LOCUS B0671259 963 bp mRNA linear EST 15-JUL-2002
DEFINITION AGNCOCRT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
IMAGE:6274716 5', mRNA sequence.
ACCESSION B0671259
VERSION B0671259.1 GI:21782093
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1 (bases 1 to 963)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LNCM2456 row: 1 column: 13  
 High quality sequence stop: 565.  
 Location/Qualifiers

## FEATURES

## SOURCE

1. 963  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1db="NIH\_MGC\_102"  
 /issue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH108 (phage-resistant)"  
 /note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 194 a 326 c 260 g 183 t

## ALIGNMENT SCORES:

Pred. No.: 2.69e-70 Length: 963  
 Score: 764.00 Matches: 151  
 Percent Similarity: 95.65% Conservative: 3  
 Best Local Similarity: 93.79% Mismatches: 7  
 Query Match: 65.75% Indels: 1  
 DB: 14 Gaps: 0

US-09-245-198a-2 (1-225) x BQ671259 (1-963)

65 VALARGPARGARGSERALAPROLYSGLYALALAPROARGARGALALEA 84  
 |||||||  
 2 GTTCGGCTCGCAGAGTGCACCTAAAGCGGAAAAA-CGGGCTCGAAGCGATCGCA 60  
 |||||||  
 85 ALALHSTYRGVALHISPRARGPROGLYGLNAPGLYALAGLVALASPGLY 104  
 |||||||  
 61 GCCCATTTATGAAGTTCAATCCACGACCTGGACAGCAGCGAGCGAGCTGGACGG 120  
 |||||||  
 105 THRVALSERGLYTPGLUGLUTHTHLYLLEANSERSESERPROLEARGTYAASPG 124  
 |||||||  
 121 ACAATGAGTGGCTGGGAGAAAGCCAGAAATCAACGCTCCACCCCTCGCGTACACCGC 180  
 |||||||  
 125 GNLLEGLGLUPHERVALILLEGALAGLYLEUTYTYTYLEUYCYGSLNVALHIS 144  
 |||||||  
 181 CAGATCGGGGAGTTATATAGTACCCGGGCTGGCTACTACTGTCAGATGTCAGTGC 240  
 |||||||  
 145 PHEASPGGLUGLYLSALVALTYTYLEULYSLEUASPLEUAEUVALASNGLYVALLEUA 164  
 |||||||  
 241 TTTGATAGGGGAGGAGCTGTCTACCTGACCTGACCTGCGATGATGCTGCTGCC 300  
 |||||||  
 165 LEUARGCYSLUGLUGLUPHESERLALATHRALALASERSEPROGLYPROGLLEUARG 184  
 |||||||  
 301 CTGCGCTGCTCGGAGAAATTTTCACGACCTCGGCGCACTTCGCGGCCACCTCCGC 360  
 |||||||  
 185 LEUCYSGNLVALISERGLYLEUAEUAEUARGPROGLYSEPSERLEARGILARGTHR 204  
 |||||||  
 361 CTCGCGCAGAGTGTCTGGGCTGTGGCCCTGGCGCCAGGCTCTCCCTCGCGATCGCAC 420  
 |||||||

QY 205 LeuProTPRALAHSLLEUYSALALAPROPhLeuThTYrPheGLYleuPhGLNVal 224  
 |||||||  
 Db 421 CTCCTCGGCGCCATTCACAGGCTGCCCTTCTCTCACCTACTTCGACCTCTCCAGGTT 480  
 |||||  
 QY 225 His 225  
 |||||  
 Db 481 CAC 483

## RESULT 11

AK020909

LOCUS

DEFINITION

Mus musculus adult retina cDNA, RIKEN full-length enriched library,  
 clone:A930030D13; tumor necrosis factor (ligand) superfamily, member  
 12, full insert sequence.

## ACCESSION

AK020909

VERSION

KEYWORDS

SOURCE

## ORGANISM

Mus musculus

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 2 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 3 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 4 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)

## TITLE

Genome Res. 10 (10), 1617-1630 (2000)

## JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

## MEDLINE

20499374

## PUBMED

11042159

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishii, K., Kikunishi, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K.,  
 Fujiwara, K., Inoue, K., Togawa, Y., Izawa, M., Ohtsuka, E., Watanabe, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multichannel sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

## JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

## MEDLINE

20530913

## PUBMED

11076861

## AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamane, I.,  
 Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fiechter, C., Gaasterland, T., Gless, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Niklido, I., Pasole, G.,  
 Quackenbush, J., Schriml, L.M., Staahl, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Balderelli, R., Barsby, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
 Toyokawa, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
 Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660







```

Db      422 COTCCCGGGGCCCATTTAAAGGCGTCCCCCTTCCTCACCATTGTGGATTTT 477
|||||
RESULT 14
LOCUS   AM917574                      584 bp    mRNA       linear     EST 25-MAY-2000
DEFINITION Rat gene index, normalized rat, norvegicus, Bento Soares
ACCESSION AM917574
VERSION   AM917574.1  GI:8083328
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 584)
AUTHORS  Lee,N.H., Gloder,A., Chandra,I., Mason,T.M., Quackenbush,J.,
          Kervaege,A.R. and Adams,M.D.
TITLE     Rat Genome Project: Generation of a Rat EST (NEST) Catalog & Rat
          Gene Index
JOURNAL   Unpublished (1998)
COMMENT   Contact: Lee, NH
          The Institute for Genomic Research
          9712, Medical Center Drive, Rockville, MD 20850, USA
          Tel.: (301)-838-3529
          Fax: (301)-838-0208
          Email: nhlee@tigr.org
          This clone is available through the ATCC, contact the ATCC
          tel#703-365-2700 for further information
          Seq primer: M13 Reverse.
FEATURES
         source                location/qualifiers
            1..584
             /organism="Rattus norvegicus"
             /db_xref="taxon:10116"
             /clone_id="RGREF49"
             /Bento_Scores="Rat gene index, normalized rat, norvegicus,
             Bento Soares"
             /tissue_type="mix - brain, ovary, placenta, kidney, lung,
             liver, embryo, heart, muscle, spleen"
             /lab_host="SOLR"
             /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
             XhoI; Estimated insert size approx.1 kb"
BASE COUNT 108 a 188 c 147 g 138 t 3 others
ORIGIN
Alignment Scores:
Pred. NO.: 9.72e-64 Length: 584
Score: 699.00 Matches: 133
Percent Similarity: 99.26% Conservative: 1
Best Local Similarity: 98.52% Mismatches: 1
Query Match: 60.15% Indels: 0
DB: 10 Gaps: 0
DS-09-245-198A-2 (1-225) x AM917574 (1-584)
QY      91 ProArgProGIyGINAspGLyAlaGlnAlaGlyValAspGLyThrValSerGIyTrpGlu 110
        |||||
Db      2 CCACAGCCAGGACAGATGAGACACAGGCAGGTGTGATGGGACAGTGAGTGGCGGA 61
        |||||
QY      111 GluThrLysIleasnSerSerSerProLeuArGIyTrnspArGInIllegIyGluPheThr 130
        |||||
Db      62 GAGACCAAAATCACAGCTCCAGCCCTCTGGCTATWACCCGCCACAATTTGGGAATTTACG 121
        |||||
QY      131 ValIleatgaIagLytleuTyrrTyrrleuTyrcysGlnValHisphasgInglyysAla 150
        |||||
Db      122 GTCAATCAGGCTGGGCTCTACTACCTGTACGTCTCAGGTGACATTTGATGAGGGAGGCA 181
        |||||
QY      151 ValTyrrleuTysIleaspLeuIleuValasnGlyValIleualaleuarGcysIleuglu 170
        |||||
Db      182 GTCTAACCGTAAGCTGGACTTGTGTGTAATGGTGTGCTGGCCCTCGCTCGCTGGAGAAGA 241
        |||||
QY      171 PheSerAlaThrAlaAlaSerSerProGIyProGInleuArGleuCysGlnValSerGIy 190

```

Dn	242	TTCACGCCACAGCAGCAGCTCTCCGTGGCCCAAGCTCCGCTGTGTGCAGGTGTCTGG	301
Oy	191	LeuteupProleuarArgProglySerSerLeuarArgileargrhrLeuproTPalaHisIeu	210
Dn	302	CTGTGCCTCTCGGCGCAGGAGTCTCCCTCGGATCGCTACCTCCCTCGGCGTCATCTT	361
Oy	211	LysAlaAlaPhePheLeuThrTyPhelyleuPhegInValHis	225
Dn	362	AAGGCTGCCCTTCCTTACTACTTTTGACTCTWTCAAGTTTCAC	406
RESULT 15			
LOCUS	AA870722	445 bp	mRNA linear EST 16-MAR-1998
DEFINITION	vq55g97.r.1 Barstead stromal cell line MFLRB8 Mus musculus cDNA		
ACCESSION	AA870722		
VERSION	AA870722.1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 445)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucada,I., Lacy,M., Le,M., Martin,D., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The MashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project MashU-HHMI Mouse EST Project Washington University School of Medicine# 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNT ; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information. MG1:601556		
FEATURES	Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 420.		
source	Location/Qualifiers 1..445 /organism="Mus musculus" /db_xref="taxon:10090" /clone_id="Barstead stromal cell line MFLRB8" /cell_line="C2C12 (undifferentiated)" /lab_host="DH10B" /note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site_1: EcoRI. Site_2: NotI. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCCTGG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Source undifferentiated tissue culture cell line C2C12, library constructed by Bob Barsstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."		
BASE COUNT	102 a	119 c	135 g 89 t
ORIGIN			
Alignment Scores:			
Pred. NO.:	7.64e-61	Length:	445
Score:	670.00	Matches:	132
Percent Similarity:	96.38%	Conservative:	1
Best Local Similarity:	95.65%	Mismatches:	5
Query Match:	57.66%	Indels:	1
DB:	9	Gaps:	0

us-09-245-198a-2 (1-225) x AA870722 (1-445)

```
OY 50 GlnThrGluGluSerGlnAspValProPheLeuGluGlnLeuValArgProArg 69
    |||
    |||
    |||
Db 12 CAAGCGTCGCAAGCGCAGGATGTGTTACTTCTTGGAACACTAGTCGGCTCGAAGA 71
    |||
OY 70 SerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAlaAlaHisTyrGluVal 89
    |||
    |||
    |||
Db 72 AGTGCTCCTAAAGGCCGGAACGG-CGGCCTCGCCGAGCTATTGCAGCCCATATGAGGTT 130
    |||
OY 90 HisProArgProGlyGlnAspGlyValGlnAlaGlyValAspGlyThrValSerGlyTyr 109
    |||
    |||
    |||
Db 131 CATCTCGGCCGACGACAGGATGAGCACAGCAGCTGTGATGGACACAGTGTGCTGG 190
    |||
OY 110 GluGluThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleGlyLuphe 129
    |||
    |||
    |||
Db 191 GAAGAGACCAAAATCACAAGCTCCAGCCCTCGCCCTACGACCGCCAGATTGGGAATTT 250
    |||
OY 130 ThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluLys 149
    |||
    |||
    |||
Db 251 ACAGTCATCAGGGCTGGGCTCTACTACTGTAAGTGTGACAGGTGCACTTGATGAGGAAAG 310
    |||
OY 150 AlaValTyrLeuLysLeuAspLeuValAsnGlyValIleuAlaLeuArgCysLeuGlu 169
    |||
    |||
    |||
Db 311 GCTGCTACCTGAAGCTGACTGTGCTGTGAAGGTGTGCTGCGCTGCGCTGCTGGAA 370
    |||
OY 170 GluPheSerAlaThrAlaAlaSerSerProGlnLeuArgLeuCysGln 187
    |||
    |||
    |||
Db 371 GAATTCACGACACAGCAGCAAGCTCTCTGGGCCCGCCAGCTCCGTTGTGTGCCAG 424
```

Search completed: May 8, 2003, 04:22:48  
Job time : 1064.98 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:11:52 ; Search time 23.8703 seconds

(without alignments)  
1942.185 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162  
Sequence: 1 VLSGLALACGLLVVSL.....PMAHLKAPFLTYEGLFQVH 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.5	9.3	325	5	09V5G2 drosophila
2	93.5	8.0	557	16	08XGX3 ralsiona s
3	90.5	7.8	210	16	09A9Z6 pseudobacter
4	90.5	7.8	421	16	09H0W2 pseudomonas
5	89.5	7.7	805	5	09VFD4 drosophila
6	89	7.7	287	13	090WT9 gallus galli
7	89	7.7	352	12	089341 hendra viru
8	88	7.6	224	5	09V762 drosophila
9	88	7.6	353	12	066760 equine morb
10	88	7.6	532	4	016727 homo sapien
11	88	7.6	1663	4	09U001 homo sapien
12	88	7.6	3261	4	09Y556 homo sapien
13	88	7.6	3664	4	096T58 homo sapien
14	87.5	7.5	522	10	09FTW7 oryza sativ
15	87.5	7.5	670	16	09AAL5 caulobacter
16	86.5	7.4	340	16	09H0R8 pseudomonas

17	86.5	7.4	504	16	092KA4 rhizobium m
18	86	7.4	260	10	08S2N9 oryza sativ
19	85.5	7.4	331	10	0942P9 oryza sativ
20	84.5	7.3	422	16	09RKB0 streptomyce
21	84.5	7.3	430	2	09REU1 streptomyce
22	84.5	7.3	1079	13	08UVR4 xenopus lae
23	84.5	7.3	1118	16	09BE34 rhizobium l
24	84	7.2	2962	5	093326 caenorhabd1
25	83.5	7.2	394	16	092V66 rhizobium m
26	83	7.1	467	16	09S2Y4 streptomyce
27	83	7.1	2779	5	09W4N7 drosophila
28	82.5	7.1	174	16	09CKX1 pasteurella
29	82.5	7.1	314	5	08WPH7 thelateria a
30	82.5	7.1	549	16	08RC38 thermoaer
31	82	7.1	619	5	08S0H9 encephalito
32	82	7.1	654	16	P74664 synecocyst
33	81.5	7.0	999	11	09JXR6 mus musculu
34	81.5	7.0	1665	11	08V111 mus musculu
35	81	7.0	258	4	09NUD0 mus musculu
36	81	7.0	262	11	09CVR1 mus musculu
37	80.5	6.9	274	10	0949I6 beta vulgar
38	80.5	6.9	324	11	09JUS4 raltus norv
39	80.5	6.9	788	4	09P2E5 homo sapien
40	80.5	6.9	1438	16	085739 pseudomonas
41	80	6.9	319	16	08U8B6 agrobacteri
42	80	6.9	403	6	09HGL8 canis famli
43	80	6.9	589	5	08TOR2 drosophila
44	80	6.9	793	5	09T108 drosophila
45	80	6.9	1522	10	08S4C3 oryza sativ

## ALIGNMENTS

RESULT 1	ID	09V5G2	PRELIMINARY:	PRT:	325 AA.
AC	09V5G2	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE	CG12919 protein.				
GN	CG12919.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BERKELEY:				
RX	MEDLINE-20196006; PubMed-10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abdill J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Batlowin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,				
RA	Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadiet E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durkin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hastin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,  
 RA Shier B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spletter E., Spredding A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003831; AAF58848.1; -  
 DR Flybase: FBgn003483; CG12919.  
 DR InterPro: IPR000478; TNF\_family.  
 DR SMART: SM00207; TNF\_1;  
 DR PROSITE: PS00251; TNF\_1;  
 DR PROSITE: PS00459; TNF\_2;  
 SQ SEQUENCE 325 AA; 36862 MW; 6E5CBB69694F1A3A CRC64;

Query Match 9.3%; Score 107.5; DB 5; Length 325;  
 Best Local Similarity 23.1%; Pred. No. 0.046; Indels 35; Gaps 9;  
 Matches 48; Conservative 34; Mismatches 91;

OY 29 QEPSEELTAEDRREPPELNPQTESQDVVPLEQLVPRRS--APKGRAPRAIAAH 86  
 DB 141 QEKSSNEATSKESPAPLHHRMRHSRH-----RHLLVKGESLLARSDSRP---AAH 191  
 OY 87 YEVRHPQDGAQAGVDGVSQMEETKINSSPLRYDQIGFTYIRAGLYLYVQVHPD 146  
 DB 192 FHLSRRRHQSGM-GYHDMYIGNDNERNISYOG-HFQTRDGYLTNTGGLYYVYQICYN 249  
 OY 147 EGRKAVYKLDLVNGVLA-----LRCLSEFSATPAASPGQLRCVSGLLPRPSS 199  
 DB 250 NSHD-----QNGFIYFGDTPPLQCLN-----TVPTNMPKAVTCHTSGILHERNER 297  
 OY 200 LRIRPL---PWAHLKADPFLYFGFLQV 224  
 DB 298 IHLKDINHNRNAVLEGGNNRSYFGIFKY 325

## RESULT 2

DB 08XOX3 PRELIMINARY; PRT; 557 AA.  
 AC 08XOX3;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Probable activation/secretion signal peptide protein.  
 GN RSP1095 OR RS02601.  
 OS *Ralstonia solanacearum* (Pseudomonas solanacearum).  
 OG *Plasmid megaplasmid*.  
 OC Bacteria; Proteobacteria; beta subdivision; *Ralstonia* group;  
 CC *Ralstonia*.  
 OX NCBI\_Taxid=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM1100;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Atilat M., Billault A., Brottier P., Camus J.C., Catalico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schliex T.,  
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissbach J., Boucher C.A.,  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646082; CAD18246.1; -

KW Plasmid; Complete proteome.  
 SQ SEQUENCE 557 AA; 60786 MW; 8B974C147D710649 CRC64;

Query Match 8.0%; Score 93.5; DB 16; Length 557;  
 Best Local Similarity 26.6%; Pred. No. 2;  
 Matches 62; Conservative 25; Mismatches 77; Indels 69; Gaps 14;

OY 2 LSLGLACGLLLVVSIGSMATLS-AQPSQELTAEDRREPPE--LNPQTESQDVV 58  
 DB 10 LATGVALA-----GLPSSWAQSPAGNPDIDLPVDTSRPEQKHVQGRPN--- 58  
 OY 59 PLEQL-----VPRRSAPGRKARRRAIAAHYEVHPRQDGAQAGVDGVSQMEETKI 114  
 DB 59 PALENLATLHPRSKRFQEGVALPPEPEIAHFA--PMAGHDVTVAOILQAAAS--EYTKL 114  
 OY 115 -NSSPLRYDRQIGFTYIRAGLYLYVQVHFDGKAVYKLDLVNGVLAQLCLEERS 172  
 DB 115 YADRGYPL-----SFAVVPQTF-----EGVVRIT--TVESGVARMRE--- 152  
 OY 173 ATAASSGP--QLRLC-----QVSGLLPRPGSSLRIRPLP 206  
 DB 153 ----GKPGLEARLRAISKHMDERPLRREPERVTGVLAQPGVITATVGP 201

## RESULT 3

DB 09A926 PRELIMINARY; PRT; 210 AA.  
 AC 09A926;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hypothetical protein CC1168.  
 GN CC1168.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;  
 CC *Caulobacter*.  
 OX NCBI\_Taxid=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Taub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Ullendack T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
 RT "Complete genome sequence of *Caulobacter crescentus*,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005794; ANR23152.1; -  
 DR TIGR: CC1168;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 210 AA; 22344 MW; 8B9830ADFBF7F45C CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 210;  
 Best Local Similarity 26.2%; Pred. No. 1.1;  
 Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

OY 2 LSLGLACGLLLVVSIGSMATLSAQPSQELTAEDRREPPELNPQTESQDVVPL 61  
 DB 32 LRLGLAVPAGVATVYQAAMKSLASKIQTQATAPLVLDKPRF--TGVLKDGRRPL 88  
 OY 62 EQLVPRRSAPGRKARRRAIAAHYEVHPRQDGAQAGV---DGTSGMEETKINSS 117  
 DB 89 ITARARBDKADQIVLTVAPLVRYGGEPRPSGATAKSGVYRAEYTLTLTDEVKITSA 148  
 OY 118 SPLRYD 123  
 DB 149 EGDFD 154

## RESULT 4

```

09HWM2 ID 09HWM2 PRELIMINARY: PRT: 421 AA.
09HWM2 AC 09HWM2:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hypothetical protein PA4851.
GN PA4851.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision: Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Stacey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004898; AAC08236.1;
KV Hypothetical protein; Complete proteome.
SQ SEQUENCE 421 AA; 46442 MW; 68C5C17099953C3E CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 421;
Best Local Similarity 22.6%; Pred. No. 2.7;
Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

OY 29 QEPSEELTAEDR-----REPELNPQTEESQDVPEL-----EQLV 65
DB 163 QEPSSGEPLAEERVSFVPLRAEDPEELRLSLDDDDDEPSGKRRHGISALDEDEESA 222
OY 66 RPRRSAPGRKARPRRAIAHYEVHPRGODGAQAGVDTGVSWEETKINSSSPPLRYDQ 125
DB 223 EQGSVAPLGSARKPDEA-----PVESLDQLRDEPLQJLAWKEPR-----KQMPRR 266
OY 126 IGEFTV---RAGLYLYCQVHDE-----GKAVYLKLDL--LVNGV 162
DB 267 LGMILLILALGCGLAQYTAHYADELARQDAYPWFQAQLCPETIGCTLPKSYDVEQLRSSN 326
OY 163 LALRCLLEESA-----TAASSRGP--QLRLCQVSGIL--PLRPSSSL 200
DB 327 LVVRSHPEFSGALVVDALITYNRASFSQPPLELRADLNGHLIANRRKPEYL 381

RESULT 5
O9VFDA ID 09VFDA PRELIMINARY: PRT: 805 AA.
O9VFDA AC 09VFDA:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG18442 protein.
GN CG18442.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

```

```

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.D., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003707; AAF55126.1;
KV Flybase; FBgn0038287; CG18442.
SQ SEQUENCE 805 AA; 86321 MW; FDF9EEEF14E69957 CRC64;

Query Match 7.7%; Score 89.5; DB 5; Length 805;
Best Local Similarity 21.7%; Pred. No. 7.7;
Matches 49; Conservative 27; Mismatches 55; Indels 95; Gaps 10;

OY 29 QEPSEELTAEDRREPELNPQTEESQDVPELEOLVPRRSAPGRKARPR--ATAAH 86
DB 482 EQGPGEEL-----QPPPTPTTEQ-----LSPPRAPPKSAELLQRRSPKQYRIAS 529
OY 87 YEYHR-----FGDGAQAGVDTGVSWEETI-----N 115
DB 530 PYWQDERRELCPOLPPRGSPPLDGSOSSTPNVAVSGPKKPLPLPIACRPRNSNGVNSPN 589
OY 116 SSS-----PLRYDROI-----GEFTVRAGLYLYCQV 143
DB 590 SSSPGSAPPRAHSPPIPAATVRPLPHLNOANGTLPPLPKKQQLHGEGLFTKNG----- 641
OY 144 HFDEGKAVYLKLDLVNGVL-----ALRCLLEFSATAASSPGPOL 183
DB 642 HLDGALLAKTDVAMSGILLIKLDVYAQC-----SVAQAAGGCTSI 683

RESULT 6
O9OWT9 ID 09OWT9 PRELIMINARY: PRT: 287 AA.
O9OWT9 AC 09OWT9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TNF-related apoptosis inducing ligand.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.

```

RA Bridgman J.T., Johnson A.L.;  
 RT "TNF-related apoptosis inducing ligand (TRAIL) expression in the hen  
 RL ovary.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY057941; AL23702.1;  
 DR InterPro; IPR003636; TNF\_abc.  
 DR InterPro; IPR00478; TNF\_family.  
 DR Pfam; PF00229; TNF; 1.  
 DR ProDom; PD002012; TNF\_abc; 1.  
 DR PROSITE; PS50049; TNF\_2; 1.  
 SO SEQUENCE 287 AA; 32092 MW; DB06E1C95087B108 CRC64;

Query Match 7.7%; Score 89; DB 13; Length 287;  
 Best Local Similarity 23.4%; Pred. No. 2.3;  
 Matches 40; Conservative 24; Mismatches 65; Indels 42; Gaps 7;

OY 84 AAHVHPR-PQDGAQAGVDG-----TWSGMEETKINS-SPLRYDRQIEFTVIRA 134  
 DB 127 SAHLFRQNPADGSSRRFGLNSQSRHATRWEDSTIHSILQNTY--RDGRLRVQDA 184  
 OY 135 GLVLYCGVHFD-----EGKAVYTKLILVNGVIALRCLEEFSAFA 175  
 DB 185 GKTYISQIYFRYSRDGAGARVSPQVQINMKTYSQPIILKGV-----GTK 234  
 OY 176 ASSPQGP--LRICQVSGILPLRPGSSLRITLPMHAKAPLTYEGLFOV 224  
 DB 225 CWAPAEVGLHALYQGGFLKAGDELFSVSLAIDYSDAASTFGAFRL 285

## RESULT 7

089341 PRELIMINARY; PRT; 352 AA.  
 AC 089341;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Matrix.  
 GN M.  
 OS Hendra virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=63330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99058172;  
 RA Yu M., Hansson E., Langedijk J.P., Eaton B.T., Wang L.F.;  
 RT "The attachment protein of Hendra virus has high structural similarity  
 RT but limited primary sequence homology compared with viruses in the  
 RT genus Paramyxovirus.";  
 RT Virology 251:227-233(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20481636; PubMed=11024125;  
 RA Wang L.F., Yu M., Hansson E., Pritchard L.I., Shiell B.,  
 RA Michalski W.P., Eaton B.T.;  
 RT "The exceptionally large genome of Hendra virus: support for creation  
 RT of a new genus within the family Paramyxoviridae.";  
 RT J. Virol. 74:9972-9979(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Wang L.F., Yu M., Pritchard L.I., Hansson E., Eaton B.T.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017149; AAC83191.1;  
 DR InterPro; IPR000982; Matrix.  
 DR Pfam; PF00661; Matrix; 1.  
 DR ProDom; PD000741; Matrix; 1.  
 SO SEQUENCE 352 AA; 39793 MW; 79E238DE496828D5 CRC64;

Query Match 7.7%; Score 89; DB 12; Length 352;  
 Best Local Similarity 28.1%; Pred. No. 3;  
 Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

DB 197 RNNNAIAFNLVLYLKIDADLAKAGIOGSFDR-DGTRKVAS-----FMLHGNF-VRRACKY 249  
 OY 138 YL--YCOVHFEDEGRKAVY-----LKLILVNGVIALRCLEEFSAFASSPQRLQCV 188  
 DB 250 YSEYCKRKRIIDMKLOFISGISTGSLHINKINGVISKRLPAQM-----GLQKNLC-- 299  
 OY 189 SGLPLRPGSSLRITLPM 207  
 DB 300 FSLMDINPM-----LNLRLTW 314

## RESULT 8

09V762 PRELIMINARY; PRT; 224 AA.  
 AC 09V762;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE CG10205 protein.  
 GN CG10205.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Ephydroidea; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Kaipen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Spier E., Spradling A.C., Stapleton M., Skupski M.P., Smith T.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003813; AAF58200.1;  
 DR FlyBase; FBgn0033970; CG10205.  
 SO SEQUENCE 224 AA; 25025 MW; 780B9EF7BBAF156 CRC64;

Query Match 7.6%; Score 88; DB 5; Length 224;



Best Local Similarity 22.6%; Pred. No. 2.1;  
Matches 44; Conservative 27; Mismatches 52; Indels 72; Gaps 10;

OY 10 CIGLLVVVSGSNATLSAQPESQOELEAED-----RREPPELNRQ--TESSQ 55  
DB 13 CIGAL-----TWASLAIEESTDIE---EDVISTANOTKQVIRVHPDLPKPDTSQ 62  
OY 56 DWPFLE-QLVVRPRS-----APGRKARPRRA-----IAHYEVHPRPGDGA 98  
DB 63 MNSPFIQVQTLRPSFAFSSQROYVDSKOMRRRPRPKLLGGNAA-----SGEDAT 115  
OY 99 QAGVDGVSGWETKINSSPLIRD-----RQIGETVIRAG-----135  
DB 116 QPALKEELKAPFKOKLKQOKQLINDARAGAGELAGROYELKQVEVGQVLPVQMTPGPY 175  
OY 136 -LYYLXQVHFDECK 149  
DB 176 PLYVVSKTNGRFGK 190

RESULT 9

O66760 PRELIMINARY; PRT; 353 AA.

ID Q66760  
AC Q66760; PRELIMINARY; PRT; 353 AA.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Matrix protein  
OS Equine morbillivirus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=45337;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96419932; PubMed=8822631;  
RA Gould A.R.;  
RT \*Comparison of the deduced matrix and fusion protein sequences of  
RT equine morbillivirus with cognate genes of the Paramyxoviridae.\*;  
RL Virus Res. 43:17-31(1996).  
DR EMBL; U49404; AAB39504.1; -.  
DR InterPro; IPR000982; Matrix.  
DR Pfam; PF00661; Matrix; 1.  
DR ProDom; PD000741; Matrix; 1.  
SQ SEQUENCE 353 AA; 40142 MW; EBE67A14B0FB2AD CRC64;

Query Match 7.6%; Score 88; DB 12; Length 353;  
Best Local Similarity 28.1%; Pred. No. 3.7;  
Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

OY 78 RPRRAIAHYEVHPRPGDGAQAGVDGVSGWETKINSSPLRYDRQIGETVIRAGLY 137  
DB 198 RRNAIAINLVYLIKIDDLAKAGIOGSFDR-DGTRKAS-----PMLHGNF-VRRAGKY 250  
OY 138 YL--YCVHFDEGKAVY-----LKDLVNGVIALRCLEFSAATASSPPQRLCOV 188  
DB 251 YSVECKRKIRIDMKLOFSLSIGSLHIKINGVSKRLFAQM-----GFOKNLC-- 300  
OY 189 SGLPLRPGSSLRITLPM 207  
DB 301 FSLMDINPM---LNRITW 315

RESULT 10

O16727 PRELIMINARY; PRT; 532 AA.

ID O16727  
AC O16727; PRELIMINARY; PRT; 532 AA.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Placental-like alkaline phosphatase precursor (EC 3.1.3.1).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE-COLON.  
RX MEDLINE=90124311; PubMed=2297757;  
RA Gum J.R., Hicks J.W., Sack T.L., Kim Y.S.;  
RT "Molecular Cloning of Complementary DNAs Encoding Alkaline Phosphatase  
RT in Human Colon Cancer Cells.";  
RL Cancer Res. 50:1085-1091(1990).  
DR EMBL; X53279; CAAG374.1; -.  
DR HSSP; P0634; IAC.  
DR InterPro; IPR001952; ALK\_phosphatase.  
DR Pfam; PF00245; alk\_phosphatase; 1.  
DR PRINTS; PR00113; ALKPHPTASE.  
DR ProDom; PD001868; ALK\_phosphatase; 1.  
DR SMART; SM00098; alkpc; 1.  
DR PROSITE; PS00123; ALKALINE\_PHOSPHATASE; 1.  
KW Signal.  
FT CHAIN 20 532 POTENTIAL.  
FT SIGNAL  
SQ SEQUENCE 532 AA; 57399 MW; 354B345CB19A0BC4 CRC64;

Query Match 7.6%; Score 88; DB 4; Length 532;  
Best Local Similarity 28.6%; Pred. No. 6.3;  
Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;

OY 76 KAPRRRAIAHYEV---HPRG---QDGAQAGVDGVSGWETKINSSPLRYDRQIGF 129  
DB 400 KARDRKA---YVLLYNGRQYVLIKDGARDVTESESGSPRYROQSAVPDGETHAGED 455  
OY 130 TVIRAGLYLYLQVHFDEGKAVYLIKDLVNGV-----LALRCLEFSA----- 173  
DB 456 VAV-----FARGPQAH-----LVHGQVQGTFIAHYMAFACLEFYTACDLAPS 498  
OY 174 ---TAASPSPQRLCOVSGSLPIRPGSSLRIT 204  
DB 499 AGTTDAHHPGPSV---VPALLPLLAGTLLILGT 528

RESULT 11

O90Q01 PRELIMINARY; PRT; 1663 AA.

ID O90Q01  
AC O90Q01; PRELIMINARY; PRT; 1663 AA.  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE KIAA0929 protein (Fragment).  
GN KIAA0929.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=99246063; PubMed=10231032;  
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosewa M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
DR EMBL; AB023146; BAA76773.1; -.  
FT NON-TER 1  
SQ SEQUENCE 1663 AA; 173871 MW; EC94D17086A8531E CRC64;

Query Match 7.6%; Score 88; DB 4; Length 1663;  
Best Local Similarity 23.1%; Pred. No. 27;  
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QBELTAEDRRREPPELNPOTESQ-----DWPFLEQLVVRPRSAPK----- 73  
DB 1400 QPRLPAGANRPPPEPHTOVORAQAEPTSPSPVSMKPD-L-PVSLPQTAPKQPLTV 1458

QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTGVSQMEETKINSSEPLRYDROICEFT 130  
 DB 1459 PTTSGSTPGGLVLPHTEROPAPKODS-----PLTSGRPVDMVQLKKTP 1505  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKDLLVNGVLAIRCLSEFSATPAASSPGQRL 185  
 DB 1506 IVMQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 1548  
 QY 186 CQ 187  
 DB 1549 AQ 1550

RESULT 12  
 QY9556 PRELIMINARY: PRT: 3261 AA.  
 AC Q9Y556;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE Hypothetical 357.0 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rhodes S., Huckle E.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL096858; CAB51072.1; -  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; rrm; 2.  
 DR SMART; SM00360; RRM; 2.  
 DR PROSITE; PS50102; RRM; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 3261 AA; 357037 MW; F8B6A645D09B66C CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3261;  
 Best Local Similarity 23.1%; Pred. No. 65;  
 Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;  
 QY 33 QEELTAEDRRPELPNPQTEESO-----DYVPLEQLVRRRSAPK----- 73  
 DB 2998 QPRLPGAPNRRPEPHTQVQRAQETGPTSPSPVSVSMKPDLPVSLPTQTAPKQPLFV 3056  
 QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTGVSQMEETKINSSEPLRYDROICEFT 130  
 DB 3057 PTTSGSTPGGLVLPHTEROPAPKODS-----PLTSGRPVDMVQLKKTP 3103  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKDLLVNGVLAIRCLSEFSATPAASSPGQRL 185  
 DB 3104 IVMQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 3146  
 QY 186 CQ 187  
 DB 3147 AQ 3148

RESULT 13  
 QY96758 PRELIMINARY: PRT: 3664 AA.  
 AC Q96758;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE Hypothetical 21, Last annotation update)  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBL\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Shi Y., Evans M., Xie W., Kao H.-Y., Orendtlich P., Tsai C.-C.,  
 RA Hon M., Downs R.M.;  
 RT "SHARP, an inducible cofactor that integrates nuclear receptor  
 repression and activation."  
 RL Genes Dev. 0:0-0(2001).  
 DR EMBL; AF356524; AAK52750.1; -  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; rrm; 4.  
 DR PROSITE; PS50102; RRM; 4.  
 KW Receptor.  
 SQ SEQUENCE 3664 AA; 402243 MW; 5228C585335EB27B CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3664;  
 Best Local Similarity 23.1%; Pred. No. 75;  
 Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;  
 QY 33 QEELTAEDRRPELPNPQTEESO-----DYVPLEQLVRRRSAPK----- 73  
 DB 3401 QPRLPGAPNRRPEPHTQVQRAQETGPTSPSPVSVSMKPDLPVSLPTQTAPKQPLFV 3459  
 QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTGVSQMEETKINSSEPLRYDROICEFT 130  
 DB 3460 PTTSGSTPGGLVLPHTEROPAPKODS-----PLTSGRPVDMVQLKKTP 3506  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKDLLVNGVLAIRCLSEFSATPAASSPGQRL 185  
 DB 3507 IVMQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 3549  
 QY 186 CQ 187  
 DB 3550 AQ 3551

RESULT 14  
 QY9FTN7 PRELIMINARY: PRT: 522 AA.  
 AC Q9FTN7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Putative Rer1 protein (Atrrera).  
 OS Oryza sativa (Rice).  
 ON Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 ON NCBL\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 clone: P0005A05."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002863; BAB16908.1; -  
 DR InterPro; IPR004932; Rer1.  
 DR Pfam; PF03248; Rer1; 1.  
 SQ SEQUENCE 522 AA; 58171 MW; AFA3EE294DFE88E7F CRC64;

Query Match 7.5%; Score 87.5; DB 10; Length 522;  
 Best Local Similarity 30.3%; Pred. No. 6.9;  
 Matches 57; Conservative 21; Mismatches 57; Indels 53; Gaps 15;  
 QY 70 SAPKGRKAPRRRAIAHVEVHP---RPG---ODGA---QAGVDGTVSGME-----ETK 113  
 DB 201 SSSKGRGEGRFAAREYVA-----SPSGFRPGFSMDGADSGTAGAAAAAAKRTDASRAFOY 256  
 QY 114 INSSPLRYDROICEFTVIRAGLYYLCOVHDE-----GKAVYLLKDLLVNGVLAIR 166  
 DB 257 LDRSTPHATGRWIGTIAV--AAIYAL--RYLVQGGIYTYTGIGITL-LNLII-GFLSPW 310  
 QY 167 CLEEFSAATASSPGQRLRCOVSGLLPLRPGSSLR--ITLP-----WHLKA---APPL 216

DB 311 VDEAFA-AASDGP-----ALPTGSDSEKFFIRRLPEKFMVAITKALIAFWM 360  
 QY 217 TYFGLEFQV 224  
 DB 361 TFEVFDV 368

## RESULT 15

Q9AA15 PRELIMINARY; PRT; 670 AA.  
 AC Q9AA15;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Alpha-glucosidase, putative.  
 GN CC0796.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Kouri H., Shetty J., Berry K.,  
 RA Usterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 RL EMBL; AE005755; AAK22781.1; .  
 DR TIGR; CC0796; .  
 KW Complete proteome.  
 SQ SEQUENCE 670 AA; 72724 MW; CBEC8FF95B9300C2 CRC64;

Query Match 7.5%; Score 87.5; DB 16; Length 670;

Best Local Similarity 24.0%; Pred. No. 9.5;

Matches 43; Conservative 24; Mismatches 75; Indels 37; Gaps 6;

QY 58 VFLEQLVLP-----RRSAPKGRKAR-----PRRAIAHYEVHPRP----- 93  
 DB 498 LPTRLILGPMOYTPGCFRNAPKDKSQFILPTVQTRGQALAMTVYDSDPLTMVSDSP 557  
 QY 94 -GODGAQAGVD---GTVSGMEETKINSSPLRYRQIGETVI--RAGLYYLQVHFDE 147  
 DB 558 ITYAASPAQIDFISAVPTSMDETRVLSC-----EIGQYIYARRKADMMVGMATNEA 610  
 QY 148 GKAVYIAKLVLVNGVIALKLEFEFSATAASSPEQRLRCOVSGLLPLRPGSSLRIRTLR 206  
 DB 611 GRTVKVPLSFLGAFSAEIRDEGAPETALKRTQSVASKSDTLTLKLAPEGGCVIRISP 669

Search completed: May 8, 2003, 02:21:23  
 Job time : 27.8703 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:47 ; Search time 8.39882 Seconds

(without alignments)  
1111.129 Million cell updates/sec

Title: US-09-245-198A-2  
Perfect score: 1162  
Sequence: 1 VLSGLALACGLLVVSL.....PMALKAAPFLTYFGLFOVH 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162	100.0	225	1	TN12_MOUSE
2	1020	87.8	249	1	TN12_HUMAN
3	109.5	9.4	272	1	TN12_CHICK
4	102.5	8.8	316	1	TN11_MOUSE
5	90.5	7.8	260	1	TN15_CANFA
6	90.5	7.8	318	1	TN11_RAT
7	88	7.6	532	1	PPBN_HUMAN
8	87	7.5	244	1	TN15_HUMAN
9	87	7.5	261	1	TN15_AOTTR
10	87	7.5	261	1	TN15_CALJA
11	86.5	7.4	240	1	TN14_HUMAN
12	86.5	7.4	1237	1	B3A2_MOUSE
13	86	7.4	261	1	TN15_HUMAN
14	86	7.4	261	1	TN15_MACMU
15	85	7.3	240	1	TN15_MACNE
16	85	7.3	1237	1	B3A2_RABIT
17	84	7.2	278	1	TN16_RAT
18	83.5	7.2	241	1	TN13_MOUSE
19	83.5	7.2	250	1	TN13_MACCEU
20	82	7.1	530	1	PPB2_HUMAN
21	82	7.1	535	1	PPB1_HUMAN
22	82	7.1	535	1	PPB3_HUMAN
23	80.5	6.9	246	1	CLOC_MOUSE
24	80.5	6.9	1234	1	B3A2_RAT
25	79.5	6.8	197	1	TN16_RABIT
26	79	6.8	920	1	PARC_SYNY3
27	79	6.8	1584	1	U104_CAEL
28	79	6.8	2468	1	MABP_HUMAN
29	78.5	6.8	260	1	TN15_FEICA
30	78.5	6.8	1164	1	PRID_ARATH
31	78	6.7	285	1	T13B_HUMAN
32	78	6.7	763	1	APP2_HUMAN
33	77.5	6.7	201	1	TN16_MACCEU

34	77.5	6.7	261	1	TN15_PIG
35	77.5	6.7	787	1	NASP_HUMAN
36	77	6.6	279	1	TN16_MOUSE
37	77	6.6	817	1	NAH1_BOVIN
38	76.5	6.6	214	1	SMP_ECOLI
39	76	6.5	695	1	APP2_MOUSE
40	75.5	6.5	250	1	TN13_HUMAN
41	75.5	6.5	575	1	MIS_BOVIN
42	75.5	6.5	999	1	OXRP_RAT
43	75.5	6.5	1058	1	UBA1_HUMAN
44	75	6.5	1004	1	SAU2_MOUSE
45	75	6.5	3305	1	APPL_MANSE

## ALIGNMENTS

RESULT	ID	SEQUENCE	STANDARD	PRT	225 AA.
AC	TN12_MOUSE	054907	09CTP2		
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (Fragment).				
GN	TNFSF12				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Peritoneal macrophage;				
RX	MEDLINE=98070415; PubMed=9405449;				
RA	Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;				
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."				
RL	J. Biol. Chem. 272:32401-32410(1997).				
RN	[2]				
RP	SEQUENCE OF 83-225 FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Retina;				
RX	MEDLINE=2108560; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G., Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustigich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:685-690(2001).				
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APP3. Weak inducer of apoptosis in some cell types. Promotes angiogenesis and the proliferation of endothelial cells. Mediates NF-kappaB activation (by similarity).				
CC	-1- SUBUNIT: Homotrimer (potential).				
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (by similarity).				
CC	-1- TISSUE SPECIFICITY: WIDELY EXPRESSED.				
CC	-1- PTM: The soluble form is produced from the membrane form by				

Proteolytic processing (BY similarity).  
 -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: AF030100; AAC53517.1; -  
 DR EMBL: AK020909; BAB32249.1; -  
 DR MGD: MGI:1196259; Tnfstf12.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 DR Cytokine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 FT NON\_TER 1  
 FT CHAIN <1 225  
 FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 70 225 MEMBER 12, MEMBRANE FORM.  
 FT TRANSMEM <1 21  
 FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT FT MEMBER 12, SECRETED FORM (BY SIMILARITY).  
 FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT FT (POTENTIAL)  
 FT FT EXTRACELLULAR (POTENTIAL).  
 FT FT SITE 69 70 CLEAVAGE (BY SIMILARITY).  
 FT DISULFD 167 186 POTENTIAL.  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 225 AA; 24781 MW; 90C412CC0480659B CRC64;  
 Query Match 100.0%; Score 1162; DB 1; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-97;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 VLSIGLALACGLLLVYVSLGSMATLSAQEPSEELTAEDRREPPELNPQTEESDYVDF 60  
 DB 1 VLSIGLALACGLLLVYVSLGSMATLSAQEPSEELTAEDRREPPELNPQTEESDYVDF 60  
 OY 61 LEQVLRPRRSAPKGRKAPRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINS SPL 120  
 DB 61 LEQVLRPRRSAPKGRKAPRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINS SPL 120  
 OY 121 RYDRQIGEFYIRAGLYLYCOVHDEKAVYKLDLVNGVLAARCLSEPSATASPG 180  
 DB 121 RYDRQIGEFYIRAGLYLYCOVHDEKAVYKLDLVNGVLAARCLSEPSATASPG 180  
 OY 121 RYDRQIGEFYIRAGLYLYCOVHDEKAVYKLDLVNGVLAARCLSEPSATASPG 180  
 DB 121 RYDRQIGEFYIRAGLYLYCOVHDEKAVYKLDLVNGVLAARCLSEPSATASPG 180  
 OY 181 POLRLQVSGLLPLRPGSSLRIRLPMVAHLKAAPLYFGGLFOVH 225  
 DB 181 POLRLQVSGLLPLRPGSSLRIRLPMVAHLKAAPLYFGGLFOVH 225  
 RESULT 2  
 TN12\_HUMAN STANDARD; PRT; 249 AA.  
 AC 043508; Q8WU27;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).  
 GN TNFSF12 OR APO3L OR DR3LG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND N-TERMINUS OF SOLUBLE FORM.  
 RC TISSUE-Tonsil, and fetal liver;  
 RX MEDLINE=98070415; PubMed=9405449;  
 RA Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,

RA Hession C., Garcia I., Browning J.L.;  
 RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that  
 RT weakly induces apoptosis.";  
 RL J. Biol. Chem. 272:32401-32410(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=fetal kidney;  
 RX MEDLINE=98228355; PubMed=9560343;  
 RA Masters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,  
 RA Ashkenazi A.;  
 RT "Identification of a ligand for the death-domain-containing receptor  
 RT APO3.";  
 RL Curr. Biol. 8:525-528(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Tonsil;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP FUNCTION.  
 RA PubMed=10085077;  
 RA Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;  
 RT "TWEAK induces angiogenesis and proliferation of endothelial cells.";  
 RL J. Biol. Chem. 274:8455-8459(1999).  
 CC -1- FUNCTION: Binds to TNFSF12/APO3. Weak  
 CC inducer of apoptosis in some cell types. Mediates NF-kappaB  
 CC activation. May promote angiogenesis and the proliferation of  
 CC endothelial cells.  
 CC -1- SUBUNIT: HOMOTRIMER (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.  
 CC skeletal muscle, brain, colon, small intestine, lung, ovary,  
 CC prostate, spleen, lymph node, appendix and peripheral blood  
 CC lymphocytes. Low expression in kidney, testis, liver, placenta,  
 CC thymus and bone marrow. Also detected in fetal kidney, liver,  
 CC lung and brain.  
 CC -1- PM: The soluble form derives from the membrane form  
 CC by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a  
 CC frameshift in position 125.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF030099; AAC1923.1; -  
 DR EMBL: AF035872; AAC39724.1; -  
 DR EMBL: BC019047; AAH19047.1; ALT\_FRAME.  
 DR Genew: HGNC:11927; TNFSF12.  
 DR MIM: 602695; -  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 DR Cytokine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 249  
 FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT FT MEMBER 12, MEMBRANE FORM.  
 FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT FT MEMBER 12, SECRETED FORM.  
 FT FT CYTOPLASMIC (POTENTIAL).  
 FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT FT (POTENTIAL).  
 FT FT EXTRACELLULAR (POTENTIAL).  
 FT FT SITE 93 94 CLEAVAGE.  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .).  
 SO SEQUENCE 249 AA; 27216 MW; E660843361C28EBA CRC64;

Query Match 87.8%; Score 1020; DB 1; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 2,6e-84;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACGLLVVSVSGSWATLSAQ-----EPQOEELTAEDRREP-----P 61  
 DB 26 LGLGLALACGLLVVSVSGSWATLSAQ-----EPQOEELTAEDRREP-----P 85  
 QY 62 EOLVPRRSGPKGRKAPRAIAHYEHPRGQDGAQGVDTVSQWEEETKINSSPPL 121  
 DB 86 NRIYVRBRAPKGRKTRARAIHAHYEHPRGQDGAQGVDTVSQWEEETKINSSPPL 145  
 QY 122 YDRQIGETVIRAGLYLYCQVHFDEGKAVYTKLDLVNGVIALRCLEEFSSATPAASP 181  
 DB 146 YNRQIGETVIRAGLYLYCQVHFDEGKAVYTKLDLVNGVIALRCLEEFSSATPAASP 205  
 QY 182 QLRICVSGILPLRPSSSLRIRTLPAHAKAPFLTYFGFLFOVH 225  
 DB 206 QLRICVSGILPLRPSSSLRIRTLPAHAKAPFLTYFGFLFOVH 249

RESULT 3  
 TNF5\_CHICK STANDARD; PRT; 272 AA.  
 AC Q918D8;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-l) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White Leghorn; TISSUE=Spleen;  
 RA Tregaskes C.A., Young J.R., Burnside J.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the absence of co-stimulus as well as IgE production in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ CC or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL: AJ243435; CAB95748.1; -  
 DR HSSP: P29965; IALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 272 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

FT CHAIN 111 272  
 FT FT  
 FT DOMAIN 1 23  
 FT TRANSMEM 24 44  
 FT FT  
 FT DOMAIN 45 272  
 FT SITE 110 111  
 FT DISULFID 190 229  
 FT CARBOHYD 124 124  
 FT CARBOHYD 146 146  
 FT CARBOHYD 251 251  
 SQ SEQUENCE 272 AA; 30862 MW; 5409F24A8E53CDD7 CRC64;

Query Match 9.4%; Score 109.5; DB 1; Length 272;  
 Best Local Similarity 22.5%; Pred. No. 0.012; Indels 53; Gaps 10;  
 Matches 58; Conservative 40; Mismatches 107;

QY 1 VLSGLALACGLLVVSVSGSWATLSAQ-----EPQOEELTAEDRREP-----P 45  
 DB 34 VQIGIVLFCVLYHMKMDKMEVLSINEDIYIFLRKQKQGTGSDKSTLDCRYLAKGR 93  
 QY 46 ELNPQTESQDVPVLEQLVPRRSPKGRK-----APRRATAHYEHPRGQDGAQ 100  
 DB 94 DLCKDRTASEELPKFEMRSHHEPHLKSRNETSVAEERKOPATHTLA-----GVKS 145  
 QY 101 GVDGTVSGWEERK-INSSPLRYDRDIGETVIRAGLYLYCQVHFDEGKA-----V 151  
 DB 146 NTVVRVLRKMTTSYAPTSLSISYH--GRIKVERKALVYIVSVSTRAASAPPTLYI 203  
 QY 152 YTKLDLVNGVIALRCLEEFSSATPAASPQOLRCOV-----SGLPLRPGSSLRIRTL 206  
 DB 204 YLYLPEDEDRRL-MKLDLDHSTSTA-----LCLEQSIREGVVELQGDVNFVNVTD 254  
 QY 207 WAHLKAPFLTYFGFLFOV 224  
 DB 255 STAVNVNPGNTYFGMFKL 272

RESULT 4  
 TN11\_MOUSE STANDARD; PRT; 316 AA.  
 ID TN11\_MOUSE  
 AC Q35235; Q35306; Q9J1Y0; Q9J1K9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 11 (Receptor activator of nuclear factor kappa B ligand) (RANKL) (TNF-related activation induced cytokine) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast differentiation factor) (ODF) (Osteoclastogenesis-inhibitory factor) (OCIF).  
 GN TNFSF11 OR RANKL OR TRANCE OR ODF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=hybridoma;  
 RA MEDLINE-97460112; PubMed-9312132;  
 RA Wong B.R., Rho J., Arron J., Robinson E., Orlinick J., Chao M., Kalachkov S., Cayan E., Bartlett F.S. III, Frankel W.N., Lee S.Y., Choi Y.;  
 RA "TRANCE is a novel ligand of the tumor necrosis factor receptor family that activates c-Jun N-terminal kinase in T cells";  
 RT J. Biol. Chem. 272:25190-25194(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Thymic lymphoma;  
 RA MEDLINE-98032977; PubMed-9367155;  
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C., Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D., Galibert L.;





















RA Gattl R.A., Derauf D.C., Belmont J.W., Conley M.E.,  
 RT "A single strand conformation polymorphism study of CD40 ligand.  
 RT Efficient mutation analysis and carrier detection for X-linked hyper  
 RT IGM syndrome.";  
 RL J. Clin. Invest. 97:196-201(1996).  
 RP [15]  
 RX VARIANTS HIGM1 ARG-36; CYS-140; SER-231; MET-254 AND GLY-227 DEL.  
 RX MEDLINE-97295077; PubMed-9150729;  
 RA Nonoyama S., Shimadzu M., Toru H., Seyama K., Ninoi H., Neubauer M.,  
 RA Yata J.-I., Ochi H.D.;  
 RT "Mutations of the CD40 ligand gene in 13 Japanese patients with  
 RT X-linked hyper-igm syndrome.";  
 RL Hum. Genet. 99:624-627(1997).  
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-  
 CC STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.  
 CC INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.  
 CC -1- SUBUNIT: HOMOTRIMER.  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN  
 CC EXTRACELLULAR SOLUBLE FORM.  
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+  
 CC T-LYMPHOCYTES.  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing.  
 CC -1- DISEASE: DEFECTS IN TNFSF5 ARE THE CAUSE OF AN X-LINKED  
 CC IMMUNODEFICIENCY WITH HYPER-IGM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE  
 CC SWITCH DEFECT CHARACTERIZED BY ELEVATED CONCENTRATIONS OF SERUM  
 CC IGM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES  
 CC PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE)  
 CC RECURRENT BACTERIAL AND OPORUNISTIC INFECTIONS, INCLUDING  
 CC PNEUMOCYSTIS CARINII PNEUMONIA AND INTRACTABLE DIARRHEA DUE TO  
 CC CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTION TREATMENT WITH  
 CC INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR,  
 CC WITH A DEATH RATE OF ABOUT 10% BEFORE ADOLESCENCE.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- DATABASE: NAME-CD40Lbase;  
 CC NOTE-European CD40L defect database (mutation db);  
 CC WWW="http://www.expasy.org/databases/cd40lbase/";  
 CC FTP="ftp://ftp.expasy.org/databases/cd40lbase/";  
 CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD154 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd154.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: X68550; CAA48554.1; -  
 DR EMBL: 215017; CAA/8737.1; -  
 DR EMBL: X67878; CAA48077.1; -  
 DR EMBL: L07414; AAA35662.1; -  
 DR EMBL: D31797; BAA06599.1; -  
 DR EMBL: D31793; BAA06599.1; JOINED.  
 DR EMBL: D31794; BAA06599.1; JOINED.  
 DR EMBL: D31795; BAA06599.1; JOINED.  
 DR EMBL: D31796; BAA06599.1; JOINED.  
 DR PIR: S25684; S25684.  
 DR PIR: S25694; S25694.  
 DR PIR: S28017; S28017.  
 DR PIR: S28852; S28852.  
 DR PIR: JH0793; JH0793.  
 DR PDB: 1ALY; 17-SEP-97.  
 DR Genew: HGNC:11935; TNFSF5.  
 DR MIM: 300386; -  
 DR MIM: 308230; -  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR004478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Antigen;  
 KW disease mutation; Polymorphism; 3d-structure.  
 FT CHAIN 1 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 Query Match 7.4%; Score 86; DB 1; Length 261;  
 Best Local Similarity 25.9%; Pred. No. 1.4;  
 Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;  
 QY 105 TVSGWEE---TKINSSPLRDVROGCEFTYRAGLYLYCQVHPEDEGRAVYIKDLLYN 160  
 DB 136 SYLQWAEKGYTWSNMLTVLENGKQ---TVKQGLYLYAQTFCNSNRDASSQAPFIAS 192  
 QY 161 GVALNCLREEF-----SATASSPPQPLRLC-----QVSGLLPLRPGSSSLRITPLPMAN 209  
 DB 193 --LCIKSPGFERILLRANANTHSARP-----CGQDSIHGGVEELPGASVFNVTDPQO 246  
 QY 210 LKAAPLYTFGLFQV 224  
 DB 247 VSHGTGFTSFGLLKL 261  
 RESULT 14  
 ID TNFS\_MACMU STANDARD; PRT; 261 AA.  
 AC 09BOC7;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-  
 DE L) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Macaca mulatta (Rhesus macaque), and  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_Taxid=9544, 9531;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND M.MULATTA VARIANTS GLN-60; VAL-204; PRO-206  
 RP AND THR-215.  
 RC SPECIES-M.mulatta, and C.torquatus atys; TISSUE-Lymphocytes;  
 RX MEDLINE-21583618; PubMed-11491535;  
 RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing and homology analysis of nonhuman primate  
 RT Fas/FasLigand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as IGE  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: AF344859; AAK37541.1; -  
 DR EMBL: AF344841; AAK37600.1; -  
 DR HSSP: P29965; 1ALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.

DR InterPro: IPR00478; TNF family.  
 DR Pfam: PF00229; TNF\_1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphism.  
 FT CHAIN 1 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 113 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 113 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT TRANSMEM 23 43 MEMBER 5, SOLUBLE FORM (BY SIMILARITY).  
 FT TRANSMEM 23 43 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 44 261 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT SITE 112 113 (POTENTIAL).  
 FT DISULFID 178 218 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 240 240 CLEAVAGE (BY SIMILARITY).  
 FT VARIANT 60 60 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARIANT 204 204 H -> Q (IN M.MILATTA).  
 FT VARIANT 206 206 I -> V (IN M.MILATTA).  
 FT VARIANT 215 215 L -> P (IN M.MILATTA).  
 FT VARIANT 215 215 A -> T (IN M.MILATTA).  
 SQ SEQUENCE 261 AA; 29366 MW; AA2E0F11C0F28ABB CRC64;  
 Query Match 7.4%; Score 86; DB 1; Length 261.  
 Best Local Similarity 25.9%; Pred. No. 14; Mismatches 24; Gaps 6;  
 Matches 35; Conservative 21; Indels 24; Gaps 6;  
 OY 105 TVSGWEE---TKINSSPLRYDROIGFTVIRAGLYLYCOVHFDEGRAVYLKLDLVN 160  
 DB 136 SVLQNAERKGYTMSNNLVLENGKQL---TVKROGLYIYAQVFCNREASQAPFIAS 192  
 OY 161 GYLARCLEEF-----SATASSPGPOLRLC-----QVSGLLPLRGSSLRIRTPMAH 209  
 DB 193 --LCJLKSPPGFERRILLRAANTHSSAKP---CGQOSIHGVEFELPGASVFNWTPDPO 246  
 OY 210 LKAAPLYFGFLFOV 224  
 DB 247 VSHGTCFTSFGLKL 261  
 RESULT 15  
 TNFS\_MACNE  
 ID TNFS\_MACNE STANDARD; PRT; 240 AA.  
 AC 09BDM7;  
 DT 15-JUN-2002 (rel. 41, Created)  
 DT 15-JUN-2002 (rel. 41, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-l) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Macaca nemestrina (Pig-tailed macaque).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9545;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymphocytes;  
 RX MEDLINE=21383618; PubMed=11491535;  
 RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing, and homology analysis of nonhuman primate  
 RT Fas/Fas-ligand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as IgE  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).

CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF344853; AAK37613.1; -  
 CC HSP: P29965; IALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR00478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 240 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 113 240 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 113 240 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT TRANSMEM 23 43 MEMBER 5, SOLUBLE FORM (BY SIMILARITY).  
 FT TRANSMEM 23 43 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 44 240 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT SITE 112 113 (POTENTIAL).  
 FT DISULFID 173 197 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 219 219 CLEAVAGE (BY SIMILARITY).  
 FT SEQUENCE 240 AA; 27273 MW; BAB6AA0213CAEDC CRC64;  
 Query Match 7.3%; Score 85; DB 1; Length 240.  
 Best Local Similarity 26.0%; Pred. No. 16; Mismatches 53; Indels 24; Gaps 6;  
 Matches 34; Conservative 20; Indels 24; Gaps 6;  
 OY 109 WEE---TKINSSPLRYDROIGFTVIRAGLYLYCOVHFDEGRAVYLKLDLVNGVLA 164  
 DB 119 WAEKGYTMSNNLVLENGKQL---TVKROGLYIYAQVFCNREASQAPFIAS--LC 173  
 OY 165 LKCLEEF-----SATASSPGPOLRLC-----QVSGLLPLRGSSLRIRTPMAHAKAA 213  
 DB 174 LKSPGFERRILLRAANTHSSKTP---CGQOSIHGVEFELPGASVFNWTPDPOVSHG 229  
 OY 214 PFLTYFGFLFOV 224  
 DB 230 TGFTSFGLKL 240

Search completed: May 8, 2003, 02:20:14  
 Job time : 10.3988 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:13:22 ; Search time 12.3772 Seconds

(Without alignments)  
1747.586 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSIGLALACILGLLVVSL.....PMAHLKAPLFTYGLFQVH 225

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.5	7.8	210	2 D87394	hypothetical prote
2	90.5	7.8	421	2 C83040	hypothetical prote
3	89	7.7	352	2 T08209	matrix protein - H
4	88	7.6	532	2 S12076	alkaline phosphata
5	87.5	7.5	670	2 A87348	alpha-glucosidase,
6	87	7.5	244	2 A46066	lymphotoxin beta -
7	86.5	7.4	340	2 A83035	probable transmem
8	86.5	7.4	1237	2 A31789	band 3-related pro
9	86	7.4	261	2 I53476	CD40 ligand - huma
10	85	7.3	1237	2 A56764	band 3-related pro
11	84	7.2	278	2 A49266	fas ligand - rat
12	84	7.2	2962	2 T19756	hypothetical prote
13	83.5	7.2	394	2 H95947	probable conserved
14	83	7.1	467	2 T34874	hypothetical prote
15	82	7.1	335	1 PAHUA	alkaline phosphata
16	82	7.1	654	2 S76870	hypothetical prote
17	80.5	6.9	246	2 S29328	complement subcomp
18	80.5	6.9	1234	2 A34911	band 3-related pro
19	80.5	6.9	1438	2 T17402	dihydroaeruginosic
20	80	6.9	261	2 AB3070	conserved hypotet
21	80	6.9	319	2 H98216	hypothetical prote
22	79.5	6.8	197	2 JH0309	tumor necrosis fac
23	79	6.8	487	2 F82065	RNA polymerase sig
24	79	6.8	944	2 S75188	DNA topoisomerase
25	79	6.8	1584	2 JN0114	kinesin-like relate
26	79	6.8	1584	2 T15822	hypothetical prote
27	78.5	6.8	289	2 H83433	hypothetical prote
28	78.5	6.8	774	2 D83507	phytochrome D - Ar
29	78.5	6.8	1164	2 B71429	

30	78.5	6.8	1396	2 A44453	translation initia
31	78	6.7	240	2 T49856	probable Alpi1 hom
32	78	6.7	590	2 AB1411	autolysin, N-acety
33	77.5	6.7	763	2 A49321	amyloid beta (A4)
34	77.5	6.7	787	2 A48819	nuclear autoantige
35	76.5	6.6	279	2 A53062	fas ligand - mouse
36	76.5	6.6	214	2 A26227	smg protein - Esch
37	76.5	6.6	562	2 A98297	probable lipote-p
38	76.5	6.6	562	2 C86138	probable lipote-p
39	76.5	6.6	1069	2 S27922	nuclear antigen EB
40	76	6.5	200	2 B83642	hypothetical prote
41	76	6.5	380	2 T34568	hypothetical prote
42	76	6.5	511	2 JC1404	CDER-box DNA-bind
43	76	6.5	751	2 A49274	beta-amyloid precu
44	75.5	6.5	466	2 T24257	hypothetical prote
45	75.5	6.5	575	1 WFBOM	mullerian inhibiti

#### ALIGNMENTS

##### RESULT 1

D87394 hypothetical protein CC1168 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: D87394

R:Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K

N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249, MUID:21173696, PMID:11259647

A:Accession: D87394

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-210 <STO>

A:Cross-references: GB:AE005673; NID:q13422488; PIDN:AMK2152.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC1168

Query Match Best Local Similarity 7.88; Score 90.5; DB 2; Length 210;

Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

QY 2 LSLGIALACILGLLVVSLGSHATISAOEPSEELTAEDRREPPPLNPQTESQDVYPL 61

DB 32 LRLGIAFVAGVAAATVIVQAAWRSIASSKLTQAAATAPLVLDKPRF--TGVLKDGRRPFL 88

QY 62 EQLVPRRSAPKGRKARRRAIAAHYEVHPRPGDGAQAGV---DGTVSGMEETKINSS 117

DB 89 ITAEAEADAKDONIVRLTAPLVLRGYPEPNSQATKSGYREAEMLTLTDEVKITISA 148

QY 118 SPLRYD 123

DB 149 EGDFPD 154

##### RESULT 2

C83040 hypothetical protein PA4851 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83040

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83040

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <STO>  
 A:Cross-references: GB:AE004898; GB:AE004091; NID:g9951115; PIDN:AA08236.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4851

Query Match 7.8%; Score 90.5; DB 2; Length 421;  
 Best Local Similarity 22.6%; Pred. No. 3.4;  
 Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

29 QEESQELTAEADR-----RPPPLNPTQESQOVVPEFL-----EQLV 65  
 DB 163 QEESQELTAEADR-----RPPPLNPTQESQOVVPEFL-----EQLV 65  
 QY 66 RPRRAAFGRKARRPRAIAHVEHPRPGDGAQAGVGTSGWEETKINSSEPLRYDRQ 125  
 DB 223 EEDSVARLGSARKRPDEA-----PVESLQDLNDEPIQLAMEKR-----RQMPRR 266  
 QY 126 IGEFTYV---RAGLYTYICQVHDE-----GKAVYTKLDL-LVNGV 162  
 DB 267 LGMULLLILALGGLAQAQYIYHFDLARDADAYRPFPAQLCPETGCTLPKRVDEQIRSSN 326  
 QY 163 LALRCLFEESA-----TAASSGP--QIRLQVSGIL---PLRPGSSL 200  
 DB 327 LVVRSHEFSFGALVVDIAITNKRASFQPFULELRFADLNGHLIARRRPFETL 381

RESULT 3  
 T08209  
 matrix protein - Hendra virus  
 C:Species: Hendra virus  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: T08209  
 R:Yu, M.; Hanson, E.; Langedijk, J.P.M.; Eaton, B.T.; Wang, L.F.  
 Virology 251, 227-233, 1998  
 A:Title: The attachment protein of hendra virus has high structural similarity but limited  
 A:Reference number: Z16405; MID:99058172; PMID:9637786  
 A:Accession: T08209  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: genomic RNA  
 A:Residues: 1-352 <YUM>  
 A:Cross-references: EMBL:AF017149; NID:g3273489; PIDN:AA083191.1; PID:g3273494  
 C:Genetics:  
 A:Gene: M  
 C:Superfamily: parampfluenza virus matrix protein  
 C:Keywords: matrix protein

Query Match 7.7%; Score 89; DB 2; Length 352;  
 Best Local Similarity 28.1%; Pred. No. 3.8;  
 Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;  
 QY 78 RPRRAIAHVEHPRPGDGAQAGVGTSGWEETKINSSEPLRYDRQIGFTYIRAGLY 137  
 DB 197 RNNNAIAFNLLVYLKIDADLAKAGIGSFDK-DGTVAS-----FMHLGNF-VRRGKY 249  
 QY 138 YL--YCVHDEGKAVY-----LKIDLNVGYLALRCLFEESATRAASSPQQLRLQV 188  
 DB 250 YSEYCYCKRIDRKRLQFSLGISIGLSLHINKINGVISKRLEFAOM-----GLQKNLC-- 299  
 QY 189 SGLPLRPGSSLRIPTLPW 207  
 DB 300 FSLMDINFW---LNRLTW 314

RESULT 4  
 S12076  
 alkaline phosphatase (EC 3.1.3.1) precursor, placental-like - human  
 N:Alternate names: alkaline phosphatase Nagao isozyme; germ-cell alkaline phosphatase  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Nov-1993 #sequence\_revision 13-Mar-1998 #text\_change 31-Jan-2000  
 C:Accession: S12076; 137991; A34333; S00974; A31336  
 R:Gum, J.R.; Hicks, J.W.; Sack, T.L.; Kim, Y.S.  
 Cancer Res. 50, 1085-1091, 1990

A:Title: Molecular cloning of complementary DNAs encoding alkaline phosphatase in hum

A:Reference number: S12076; MID:90124311; PMID:2297757

A:Accession: S12076

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-177, 'A', 179-272, 'W', 274-497, 'S', 499-530, 'T', 532 <GUN>

A:Cross-references: EMBL:X53279; NID:g28624; PIDN:CAA37374.1; PID:g28625

R:Lower, M.E.; Straus, A.W.

Cancer Res. 50, 3956-3962, 1990

A:Title: Expression of a Nagao-type, phosphatidylinositol-glycan anchored alkaline ph

A:Reference number: 137991; MID:90283879; PMID:2162249

A:Accession: 137991

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-151, 'V', 153-532 <LOW>

A:Cross-references: EMBL:X55958; NID:g296797; PIDN:CAA39425.1; PID:g296798

R:Matnabe, S.; Matnabe, T.; Li, W.B.; Soong, B.W.; Chou, J.Y.

J. Biol. Chem. 264, 12611-12619, 1989

A:Title: Expression of the germ cell alkaline phosphatase gene in human choriocarcino

A:Reference number: A34333; MID:89308696; PMID:2745460

A:Accession: A34333

A:Molecule type: mRNA

A:Residues: 1-379, 'L', 381-497, 'R', 499-532 <MAT>

A:Cross-references: GB:J04948; NID:g178418

R:Shen, L.P.; Liu, H.; Kan, Y.W.; Kam, W.

Nucleic Acids Res. 16, 5694, 1988

A:Title: 5' nucleotide sequence of a putative human placental alkaline phosphatase-11

A:Reference number: S00974; MID:88262578; PMID:3387245

A:Accession: S00974

A:Molecule type: DNA

A:Residues: 1-157 <SHE>

A:Cross-references: EMBL:X07247; NID:g35509

A:Note: The authors translated the codon CCG for residue 28 as Cys and AGT for residu

R:Millan, J.L.; Manes, T.

Proc. Natl. Acad. Sci. U.S.A. 85, 3024-3028, 1988

A:Title: Semima-derived Nagao isozyme is encoded by a germ-cell alkaline phosphatas

A:Reference number: A31336; MID:88203632; PMID:2834730

A:Accession: A31336

A:Molecule type: DNA

A:Residues: 1-56, 'W', 58-151, 'V', 153-315, 'L', 317-532 <MIT>

A:Cross-references: GB:J03252; NID:g178427; PIDN:AAA98616.1; PID:g178428; GB:J03512

C:Genetics: 23/1; 62/1; 100/3; 159/1; 216/3; 261/3; 286/1; 331/1; 395/1; 431/1

A:Note: the first intron occurs before the initiator codon

C:Function:

A:Description: catalyzes the hydrolysis of phosphate monoesters

C:Superfamily: alkaline phosphatase

C:Keywords: blocked carboxyl end; glycoprotein; homodimer; lipoprotein; membrane prot

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-503/Product: alkaline phosphatase, placental-like #status predicted <MAT>

F:504-532/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:111/Active site: Ser (phosphoserine intermediate) #status predicted

F:141,268/Binding site: carboxylate (Asn) (covalent) #status predicted

F:185/Binding site: substrate phosphate (Arg) #status predicted

F:339,379,451/Binding site: zinc (His) #status predicted

F:503/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asp) (in mature f

Query Match 7.6%; Score 88; DB 2; Length 532;  
 Best Local Similarity 28.6%; Pred. No. 7.3;  
 Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;

QY 76 KARRRAIAHVEV---HPRPG---ODGAQAGVGTSGWEETKINSSEPLRYDRQIGF 129  
 DB 400 KARDRA---YTVLLXNGPGVGLDGAAPDVTESESGSPVYRQOSAVPLDDETHAGD 455  
 QY 130 TVRAGLYTYICQVHDEGKAVYTKLDLVNGV-----LALRCLFEESA----- 173  
 DB 456 VAV-----FARGQAH-----LVHGQDQFTAHVMAFAACLEPYACDLAP 498  
 QY 174 ---TAASSPQQLRLQVSGILPLRPGSSLRIPT 204  
 DB 499 AGTTDAHHPSPV---VPALPLLAGTLLLTGT 528

## RESULT 5

A87348

alpha-glucosidase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: A87348

R.Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-670 &lt;STO&gt;

A:Cross-references: GB:AE005673; NID:g13422039; PIDN:AAK22781.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0796

Query Match

Best Local Similarity

Matches

43; Conservative

24; Mismatches

75; Indels

37; Gaps

6;

58 VFLEQLVLP-----RRSAPKGRKAR-----PRRAIAHYEHVHRP-----93

498 LPFTMLGPMIDYTPGGRNNAKPKFSQFIPTVQTRGQALMAYVDSPLTMSVSSP 557

OY 94 -GQDAQAGVD---GTVSGMEETKINSPLRYDRQIGFTYI--RAGLYLYCOVHEDP 147

DB 558 IYYAASPAGLDRIASVPTSMDETIVLSG-----EIGQYIYARRKKADMMVGMATNNA 610

OY 148 GAAVYLKLDLVNGLALRCLLEFSATASSPGQRLCQVSGGLPLRPGSSLRITPL 206

DB 611 GRTVVPISFIGNGAFSAEIRREDGAEPALTKRTQSVASKOTLTKLAPGGGIVIRISP 669

RESULT 6

A46066

Lymphotoxin beta - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C:Accession: A46066

R.Browning, J.L.; Ngam-ek, A.; Lawton, P.; Demarinis, J.; Tilzard, R.; Chow, E.P.; Hessio

Cell 72, 847-856, 1993

A:Title: Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric com

A:Reference number: A46066; MUID:93208881; PMID:791655

A:Accession: A46066

A:Status: preliminary

A:Molecule type: DNA; protein

A:Residues: 1-244 &lt;BRO&gt;

A:Cross-references: GB:L11015; NID:g292276; PIDN:AAA36191.1; PID:g292277

A:Note: sequence extracted from NCBI backbone (NCBIN:128066, NCBIPI:128067)

C:Keywords: transmembrane protein

Query Match

Best Local Similarity

Matches

64; Conservative

26; Mismatches

80; Indels

108; Gaps

13;

OY 3 SIGLALA---CLGLLLVNVSLGSMATLSAQSPQELTAEDR-----REPEL 47

DB 18 SLLAVAGTSLVTLAVPIVLAIVPDDGGVLTETDPGAGQAGGFGPKLPEE 77

OY 48 NPQTESQDVVPELQVLRPRSAKGRKARRAIAHYEHVPRPGDGAQAGVDGTVS 107

DB 78 EPEET---DLSPGLP-----AAHLIGAPLKGGQ-----L 102

OY 108 GMEETKINS--SSPLRYDRQIGFTYIAGLYLYCOVH-----DEGKAVYLKL 155

DB 103 GMEETKQDAFLTSGTQSPDAG--LALPDGGLYLYCLVGYGRADPPGGDPPGRSVTLKS 161

OY 156 DLVNGVIALRCLLEFSATASSPG--POLRL-----COVS 189

DB 162 SL-----YRAGGAYGCGTPELLLGCAFTYVYLDPARRGVGLMTYSVGRG 208

OY 190 GLPLRPSSSLRIRLPLMAHLKAAPL---TYFGLEPV 224

DB 209 GLVQLRRGERVYVNI---SHPMVDFARKGKTFEGAVMV 243

## RESULT 7

A83035

probable transmembrane sensor PA4895 [imported] - Pseudomonas aeruginosa (strain PAC

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 15-Jun-2001

C:Accession: A83035

R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, R.R.; Kas, A.; Lardig, R.;

Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83035

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 &lt;STO&gt;

A:Cross-references: GB:AE004902; GB:AE004091; NID:g9951162; PIDN:AG08280.1; GSPDB:G

C:Genetics:

A:Gene: PA4895

C:Superfamily: Pseudomonas putida regulatory protein pupr

Query Match

Best Local Similarity

Matches

67; Conservative

26; Mismatches

85; Indels

93; Gaps

14;

5 GLIACGLLLV-----VSLGSMATLSAQSPQEL-----36

DB 104 GLIALLGAPLVLMGMRQAQVWRDGFCAVLTAVGRRDLVEDSGOVEMTDSALDVRYD 163

OY 37 -----TAEDEPELNPQTESQDVPEL-----EQVLRPRSAKGRK 76

DB 164 AGORILLRYGEIYRTADREPR-----PELVTRDQGLRALGTFVSRR 211

OY 77 APRRAIAHYE--VHPRPGDGAQAGVDGTVSGMEETKINSPLRYDRQIGFTYI 134

DB 212 BEAEVILAV-YEGANQVPR--EGASADGR-----IEGQVRRDRQ-----RI 254

OY 135 GLIYLYCOVHEDGKAVYLIKLDLVNGLALRCLLEFSATASSPGQRLCQVSGGLPL 194

DB 255 G-----PVESASFAALMRQGLLVADDPPLR---QMGELMRVYGESIE-CEPS-LDPL 303

OY 195 RPSGLRITPLPMALKAAPLITVGLFOVH 225

DB 304 RVSGTFPPVDLP---LALAMLAQTHGLRLVH 331

RESULT 8

A31789

band 3-related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 20-Aug-1999

C:Accession: A31789

R.Alpert, S.L.; Kopito, R.R.; Libresco, S.M.; Lodish, H.F.

J. Biol. Chem. 263, 17092-17099, 1988

A:Title: Cloning and characterization of a murine band 3-related cDNA from kidney an

A:Reference number: A31789; MUID:89034212; PMID:3182834

A:Accession: A31789

A:Molecule type: mRNA

A:Residues: 1-1237 &lt;ALP&gt;

A:Cross-references: GB:J04036; NID:g192132; PIDN:AAA65505.1; PID:g309114

C:Superfamily: band 3 anion transport protein

C:Keywords: transmembrane protein

Query Match

Best Local Similarity

Matches

7.4%; Score 86.5; DB 2; Length 1237;





Best Local Similarity 22.9%; Pred. No. 7.8;  
Matches 56; Conservative 30; Mismatches 67; Indels 92; Gaps 15;

OY 15 LVVVSIGSNATLSAQEPSEELTAEDRREPELNPOTEESSQDVVPELQVLRPRNSAPKG 74  
DB 91 LVGMSIGMQLPFLD-----KELAEIREFTHNSLVSSEKQIANP--STPSE 136  
OY 75 RKARPRAIAHYEVHPRPGDGAQAGVGTGSGMEETKINS-SSPLARDROIGETVTR 133  
DB 137 TK-KPRS--VAHLTGPNRSTRIPLE-----WEDYGTALISGVTKK--GGIIVNE 182  
OY 134 AGLVLYLCOVHFDKRA-----VYLKLDLVNGVIALRCLEFSATASSPQQR 184  
DB 183 AGLVLYSKVYF-RQGSCHSOPLSHKVYMR-NPKYRGDLVL-MER-----KKLN 228  
OY 185 LCOVSGLPRLRPGSSLRIRTLPMWAh-----LKAPFL-----TYF 219  
DB 229 YCTTGOI-----WAHSSYLGAVERNLTVAADHLVYNISQSLINFEESKTFE 273  
OY 220 GLEOV 224  
DB 274 GLTKL 278

## RESULT 12

hypothetical protein C35C5.1 - *Caenorhabditis elegans*

C.Species: *Caenorhabditis elegans*

C.Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C.Accession: T19756; T21561

R.White, S.

submitted to the EMBL Data Library, August 1996

A.Reference number: T19173

A.Accession: T19756

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-2962 <W12>

A.Cross-references: EMBL:278417; PIDN:CA01693.1; GSPDB:GN00028; CESP:C35C5.1

A.Experimental source: clone C35C5

R.Harris, B.

submitted to the EMBL Data Library, August 1996

A.Reference number: T19441

A.Accession: T21561

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-2962 <W12>

A.Cross-references: EMBL:278543; PIDN:CA01756.1; GSPDB:GN00028; CESP:C35C5.1

A.Experimental source: clone F2966

C.Genetics:

A.Gene: CESP:C35C5.1

A.Map position: X

A.intons: 373/3; 452/2; 525/2; 576/3; 598/3; 782/2; 894/1; 937/3; 1334/3; 1360/3; 1587/

Query Match

Best Local Similarity 21.6%; Pred. No. 1.1e+02;

Matches 42; Conservative 29; Mismatches 91; Indels 32; Gaps 7;

OY 29 QEPSEELTAEDRREPELNPOTEESSQDVV-----PFLQVLRPRNSAPK-----GRKA 77  
DB 128 EQLLEHLNAYRHEEEEC---FEEDDITELPLPVKPAVKPRKRLPHLSTESGSTA 184  
OY 78 RPRRAIAHYEVHPRPGDGAQAGV--DGTSGMEETKI-----NSSPLARDROIGER 129  
DB 185 KSKSLVAEYVHPRPVNRMKRAVYDDGKVVQKRTYTRNIMSTIIQYHLEGETER 244  
OY 130 TVIRAGLYLYCOVHFDKRAVYLYKLDLVNGVIALRCLEFSATASSPQQLRCQVS 189  
DB 245 GRVKESTLSKTEQYLQAKIVSPKCD-----QFREQIVATAVEYDG--SVKMLQFE 294  
OY 190 GLLPLRPGSSLRIR 203  
DB 295 NALKRHSQKORL 308

## RESULT 13

H95947

probable conserved membrane-anchored protein SMD21182 [imported] - *Sinorhizobium mel*

C.Species: *Sinorhizobium meliloti*

C.Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C.Accession: H95947

R.Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; He

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A.Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing

A.Reference number: A95842; MUID:21396508; PMID:11481431

A.Accession: H95947

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-394 <KUP>

A.Cross-references: GB:AL591985; PIDN:CAC94248.1; PID:G15140734; GSPDB:GN00167

A.Experimental source: strain 1021, megaplasmid pSymB

R.Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A.Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh

A.Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A.Reference number: A96039; MUID:21368234; PMID:11474104

A.Contents: annotation

C.Genetics:

A.Gene: SMD21182

A.Genome: plasmid

C.Superfamily: Mycobacterium tuberculosis hypothetical protein Rv3272

Query Match 7.2%; Score 83.5; DB 2; Length 394;

Best Local Similarity 24.3%; Pred. No. 13;

Matches 63; Conservative 25; Mismatches 92; Indels 79; Gaps 12;

OY 13 LLLVVSIGSW-----ATLSAQEPSEELTAEDRRE--PELNPOTEESSQDVVPELQV 65  
DB 11 LELARILAPWIGOTLADGADIVKVESPADDTFTWGPPEVGEDEERLDAA-YRHACN 69  
OY 66 RPRRS-----APKGRKAPRRAIAA-----HYE-----VHRP----- 93  
DB 70 RGRRSVLDFTTAEGEAVRRLAAGSDVLENEKVGGLAKYGDYBSLSKVNRLIYCSV 129  
OY 94 ---GODGA---QGVGVGTSGMEETKINSPLRYROIG-EFTVIRAGLY-----YL 139  
DB 130 TGGGQDGPYAHRAIGYIYQMGSGIMDLTGEPRERQKIVAPADITGLTYGVIAQAL 189  
OY 140 YCOVHDEKRAVYLYKLDLVNGVIALRCLEFSATASSPQQLRCQVSGLLPLRPGSS 199  
DB 190 AGERTEGEGQIDMALDCMTGVLANQALNFL-----VSGKAPRRLGN- 232  
OY 200 LRIRTLPMWAKRAAPFLTY 218  
DB 233 -----AHPNIAPYOVF 243  
RESULT 14  
T34874  
hypothetical protein SC3A3.09 - *Streptomyces coelicolor*  
C.Species: *Streptomyces coelicolor*  
C.Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C.Accession: T34874  
R.Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M  
submitted to the EMBL Data Library, August 1999  
A.Reference number: T21560  
A.Accession: T34874  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-467 <SEB>  
A.Cross-references: EMBL:AL109849; PIDN:CA52866.1; GSPDB:GN00070; SCOEDB:SC3A3.09  
A.Experimental source: strain A3(2)  
C.Genetics:  
A.Gene: SCOEDB:SC3A3.09

Query Match 7.1%; Score 83; DB 2; Length 467;  
 Best Local Similarity 23.6%; Pred. No. 17;  
 Matches 57; Conservative 22; Mismatches 91; Indels 72; Gaps 11;

4 LGALAC-----GILLVVSIGSM---ATLSAQPSQELTAE--DKREPP-- 45  
 112 LGRVAGCLGKRPVEKPLRPLGRLGGRAGNMPPTAYFARQVPRDLTAHWNRRSAPTS 171  
 46 ---ELNPTESQDVVPLEQV---RPRRSAPKGRKAR-----PRAIA----- 85  
 172 LAENIDGPREDDNLPLGLGILLRRHCKAFTTGDAVWLDLPRGTFPAERVALNLT 231  
 86 -----HYEVHPRPGDGAQAGVGTVSWEETKINSPLRDROIGFTVIRAGLYX 138  
 232 LTGLEPHTARHNPFRFEMIGALIRADVHGW---TNGDP-----GAIAE----- 273  
 139 LYCOVHPEBGKAVYIKLDLVNGLALRCLKEFSSTAASFGPQLRCLCVGLPLRPGS 198  
 274 ---QAHND-----ATLSHTANGVHAAMFAAAVVAAPDDAPDVHACLRAGLAVVPPGS 324

OY 199 SL 200  
 Db 325 RL 326

RESULT 15  
 PAM0A  
 alkaline phosphatase (EC 3.1.3.1) precursor, placental (validated) - human  
 N:Alternate names: glycerophosphatase; orthophosphoric monoester phosphohydrolase, Regan  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1987 #sequence, revision 28-Jul-1995 #text, change 08-Dec-2000  
 C:Accession: A31074; A27363; A24318; B24318; A25385; A94099; A94082; A28958; A90114; A24  
 R:Knoll, B.J.; Rothblum, K.N.; Longley, M.  
 J. Biol. Chem. 263, 12020-12027, 1988  
 A:Title: Nucleotide sequence of the human placental alkaline phosphatase gene. Evolution  
 A:Reference number: A31074; MUID:8829886; PMID:3042787  
 A:Accession: A31074  
 A:Molecule type: DNA  
 A:Residues: 1-535 <KNO>  
 A:Cross-references: GB:J03931; GB:M19159; NID:g178475; PIDN:AAA51710.1; PID:g178476  
 R:Knoll, B.J.; Rothblum, K.N.; Longley, M.  
 Gene 60, 267-276, 1987  
 A:Title: Two gene duplication events in the evolution of the human heat-stable alkaline  
 A:Reference number: A91586; MUID:86167830; PMID:3443302  
 A:Accession: A27363  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <KN2>  
 A:Cross-references: GB:M19160  
 R:Henthorn, P.S.; Knoll, B.J.; Raducha, M.; Rothblum, K.N.; Slaughter, C.; Weiss, M.; La  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5597-5601, 1986  
 A:Title: Products of two common alleles at the locus for human placental alkaline phosph  
 A:Reference number: A94113; MUID:86287303; PMID:3461452  
 A:Accession: A24318  
 A:Molecule type: mRNA  
 A:Residues: 23-535 <HEN>  
 A:Cross-references: GB:M14169  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 A:Note: this is polymorphism type 1; 67-Leu also found  
 A:Accession: B24318  
 A:Molecule type: mRNA  
 A:Residues: 1-24, 'L', 26-65, 'V', 67-262, 'H', 264-276, 'R', 278-284, 'A', 286-388, 'C', 390-393, 'G  
 A:Cross-references: GB:M4170; NID:g178469; PIDN:AAA51709.1; PID:g178470  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 A:Note: this is polymorphism type 3; 89-Leu also found  
 R:Millan, J.L.  
 J. Biol. Chem. 261, 3112-3115, 1986  
 A:Title: Molecular cloning and sequence analysis of human placental alkaline phosphatase  
 A:Reference number: A25385; MUID:86140079; PMID:3512548  
 A:Accession: A25385  
 A:Molecule type: mRNA  
 A:Residues: 1-230, 'P', 232-535 <MIL>  
 A:Cross-references: GB:M13077; NID:g178471; PIDN:ACC97139.1; PID:g178474

R:Ovitt, C.E.; Strauss, A.W.; Alpers, D.H.; Chou, J.Y.; Boine, I.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 3781-3785, 1986  
 A:Title: Expression of different-sized placental alkaline phosphatase mRNAs in plac  
 A:Reference number: A94099; MUID:86233318; PMID:3459156  
 A:Accession: A94099  
 A:Molecule type: mRNA  
 A:Residues: 382-435, 'T', 437-535 <OVI>  
 A:Note: a soluble or serum form was also found that may arise by proteolytic cleavag  
 R:Kam, W.; Clauser, E.; Kim, Y.S.; Kan, Y.W.; Rutter, W.J.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 8715-8719, 1985  
 A:Title: Cloning, sequencing, and chromosomal localization of human term placental al  
 A:Reference number: A94082; MUID:86094295; PMID:3001717  
 A:Accession: A94082  
 A:Molecule type: mRNA  
 A:Residues: 6-24, 'L', 26-260, 'GE', 263-323, 'H', 325-395, 'FI', 398-400, 'A', 402-535 <KAM>  
 A:Cross-references: GB:M1251; NID:g178463; PIDN:AAA51706.1; PID:g178464  
 R:McAnovic, R.; Bailey, C.A.; Brink, L.; Gerber, L.; Pan, Y.C.E.; Hulmes, J.D.; Ude  
 Proc. Natl. Acad. Sci. U.S.A. 85, 1398-1402, 1988  
 A:Title: Aspartic acid-484 of nascent placental alkaline phosphatase condenses with  
 A:Reference number: A28958; MUID:88144444; PMID:3422741  
 A:Accession: A28958  
 A:Molecule type: Protein  
 A:Residues: 485-535 <MIC>  
 R:Ezra, E.; Blacher, R.; Udenfriend, S.  
 Biochem. Biophys. Res. Commun. 116, 1076-1083, 1983  
 A:Title: Purification and partial sequencing of human placental alkaline phosphatase  
 A:Reference number: A90114; MUID:84079906; PMID:6651840  
 A:Accession: A90114  
 A:Molecule type: Protein  
 A:Residues: 22-64 <EKR>  
 C:Comment: This is the form of the enzyme predominantly expressed in the placenta aft  
 C:Genetics:  
 A:Gene: GDB:ALPP; ALP1  
 A:Cross-references: GDB:119672; OMIM:171800  
 A:Map position: 2q37.1-2q37.1  
 A:Introns: 26/1; 65/1; 103/3; 162/1; 219/3; 264/3; 289/1; 334/1; 398/1; 437/1  
 A:Note: this gene is very polymorphic  
 C:Complex: homodimer  
 C:Function:  
 A:Description: catalyzes the hydrolysis of phosphate monoesters  
 C:Superfamily: alkaline phosphatase  
 C:Keywords: blocked carboxyl end; glycoprotein; homodimer; lipoprotein; membrane prot  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-506/Product: alkaline phosphatase, placental #status experimental <MAT>  
 F:507-535/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:114/Active site: Ser (phosphoserine intermediate) #status predicted  
 F:144,271/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:188/Binding site: substrate phosphate (Arg) #status predicted  
 F:342,382,454/Binding site: zinc (His) #status predicted  
 F:506/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asp) (in mature t

Query Match 7.1%; Score 82; DB 1; Length 535;  
 Best Local Similarity 27.9%; Pred. No. 24;  
 Matches 43; Conservative 15; Mismatches 45; Indels 50; Gaps 8;

OY 76 KAPRRRAIAHYEV---HPRPG---QDGAQGVGTVSWEETKINSPLRDROIGEF 129  
 403 KARRRKA---YTVLLYNGNGGVYLDKGDARDYTESGSEYRQOSAVVLDERTHAGED 458  
 OY 130 TVIRAGLYLYLCOVHPEBGKAVYIKLDLVNGLVNGV-----LALCLESFSA----- 173  
 459 VAV-----FARGPOAR-----LVHNGVEQTFIAHVMFAACLEPTACDLAPP 501  
 OY 174 ---TAASPGPOLRLCOVNSGLPLRPSSLRIRT 204  
 502 AGTTDAHPRGSV---VPALPLLAAGTLILLER 531

Search completed: May 8, 2003, 02:22:03  
 Job time : 15.3772 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

## SUMMARIES

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:47 ; Search time 166.65 Seconds

(without alignments)  
3040.496 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSIGLALACTGLLVVSL.....PMAHLKAPFLTYGFLQVH 225

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xip  
-Q=/cgr2\_1/USFTO.spool/US09245198/rnat\_06052003\_170239\_9605/app.query.fasta\_1.846  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US09245198.ecgn\_1.1.323 @rnat\_06052003\_170239\_9605 -NCPU=6 -ICPU=3  
-NO\_XLPRY -NO\_MMAR -LARGEQUERY -NBS\_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	1162	100.0	1168	19	AAV18599	Mus musculus tumour
2	1089	93.7	701	20	AAV18599	Mouse TRNL3 DNA.
3	1020	87.8	1236	19	AAV17613	TNF related endoth
4	1020	87.8	1236	21	AAV17613	Human TREPA (TNF r
5	1020	87.8	1353	22	AAV49717	Human PRO207 CDNA
6	1020	87.8	1353	24	ABK40255	CDNA encoding huma
7	1020	87.8	1353	24	ABK40255	Human CDNA encodin
8	1020	87.8	1373	19	AAV18600	Homo sapiens tumou
9	1020	87.8	1421	20	AAV56000	Human tumour necro
10	951	81.8	898	22	AAV50964	Expression vector
11	951	81.8	1030	20	AAV23424	Human TRNL3 DNA.
12	113.5	9.8	2148	24	ABK11680	DNA encoding tumou
13	111	9.6	1221	24	ABK11679	DNA encoding tumou
14	110.5	9.5	978	23	ABK11473	Drosophila melanog
15	106.5	9.2	1630	19	AAV41377	NE-KB receptor act
16	106.5	9.2	1630	19	AAV41371	NE-KB receptor act
17	106.5	9.2	1630	22	AAV15310	Murine receptor ac
18	106.5	9.2	1630	22	AAV08714	Murine RANKL (re
19	106.5	9.2	1630	22	AAV05903	Human secreted pro
20	106.5	9.2	3878	22	AAV05230	Human secreted pro
21	106.5	9.2	4034	22	AAV54040	Human prostate CDNA
22	106.5	9.2	4034	22	AAV53868	Human prostate CDNA
23	106.5	9.2	4034	24	ABK195411	Human P53S splice
24	106.5	9.2	4894	22	AAV54038	Human prostate CDNA
25	106.5	9.2	4894	22	AAV53866	Human P53S splice
26	106.5	9.2	4894	24	ABK195409	Human P53S splice
27	106.5	9.2	6976	22	AAV54041	Human prostate CDNA
28	106.5	9.2	6976	22	AAV53869	Human P53S splice
29	106.5	9.2	6976	24	ABK195412	Human P53S splice
30	105.5	9.1	81800	22	ABK84756	Human CDNA differe
31	103	8.9	684	22	AAV82935	HIV-1 gp120 V3 loo
32	102.5	8.8	951	19	AAV69900	Nucleotide sequenc
33	102.5	8.8	951	21	AAV39156	Mouse OSM nucleoti
34	102.5	8.8	951	21	AAV39965	DNA encoding a mur
35	102.5	8.8	951	21	AAV49024	Nucleic acid encod
36	102.5	8.8	1538	19	AAV69886	Osteoclast formati
37	102.5	8.8	2166	24	ABK11678	DNA encoding tumou
38	102.5	8.8	2191	19	AAV41489	Nucleotide sequenc
39	102.5	8.8	2237	20	AAV80224	Murine TRANCE enco
40	102.5	8.8	2237	24	ABK12880	CDNA encoding mous
41	102.5	8.8	2295	19	AAV70284	Human osteoprotege
42	102.5	8.8	2299	21	AAV29966	DNA encoding a mur
43	101.5	8.7	7332	24	ABK40768	Constitutively act
44	99.5	8.6	38186	20	AAV23028	Human MTH1 relate
45	99.5	8.6	38186	22	AAV30085	AC004449 CDNA clon

## ALIGNMENTS

RESULT 1  
AAV18599 standard; CDNA; 1168 BP.

AAV18599;  
21-JUL-1998 (first entry)

Mus musculus tumour necrosis factor related ligand (TRELL) gene.

TRELL: tumour necrosis factor related ligand; tnfr; treatment;

KW cancer; autoimmune disease; immune system; stimulation; suppression;

KW graft rejection; ds.

XX Mus musculus.

OS Location/Qualifiers

XX Key 2..679

FT CDS



developmental or gestational abnormalities

XX Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family

CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active

CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or

CC their active fragments. ApO4 is useful for diagnosing prostate cancer

CC by determining levels of ApO4 in an individual. Prostate cancer can also

CC be treated using ApO4 selective binding agents linked to a therapeutic

CC moiety. ApO4 polypeptides are also useful for identifying selective

CC binding agents, useful in diagnosis/treatment of disease by binding of

CC agents to the polypeptide/active fragment which is extracellular, or

CC expressed on the cell surface. The binding is preferably performed in

CC vivo. ApO4 polypeptides/active fragments are also useful for screening

CC for agonists and antagonists by binding and observing the change in ApO4

CC activity. Effective pharmacological agents useful in diagnosis or

CC treatment of disease are also identified using ApO4 polypeptides/active

CC fragments and ApO4 signal transducer molecules that specifically interact

CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4

CC activity. The method is performed in vivo or in vitro. ApO polypeptides

CC are all useful as immunogens for preparing antibodies. ApO4 is also

CC useful for diagnosis/treatment of developmental or gestational

CC abnormalities. ApO8 was transfected to human breast carcinoma cell line

CC MCF-7, and induced apoptosis.

XX

SQ Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	1,94e-90	Length:	701
Score:	1089.00	Matches:	210
Percent Similarity:	99.53%	Conservative:	0
Best Local Similarity:	99.53%	Mismatches:	1
Query Match:	93.72%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-2 (1-225) x AAX23425 (1-701)

QY 15 LeuValValValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 34

DB 1 CTGGTCTGGTGCACCTGGGAGAGCTGGGCAACGCTGTCTGCCAGAGCTTCTCAGAG 60

QY 35 GluLeuThrAlaGlnAlaSPARGLUPROProGluLeuAsnProGlnThrGlnGluSer 54

DB 61 GAGCTGACAGCAGAGAGAGCCGGGAGGCCCTGACAGCATCCCAACAGAGAGAAAGC 120

QY 55 GlnAspValValProPheLeuGlnGlnLeuValAlaArgProArgSerAlaProGly 74

DB 121 CAGGATGTGTACTTCTTCTGGAACAACCTAGTCCGGCTCGAAGAGTCTCTAAAGGC 180

QY 75 ArgValAlaArgProArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 94

DB 181 CGGAGGCGGCGGCTCGGCGAGCTATGTCACCCCAATATGAGGTTCATCTCGGCGAGA 240

QY 95 GlnAspGlnAlaGlnAlaGlnValAlaSPolYThrValSerGlyTTPGlnGluThrLysIle 114

DB 241 CAGATGGAGCACAAGAGAGGTGTGATGGAGAGAGAGTGGCTGGGAAGACCAAAATC 300

QY 115 AsnSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAla 134

DB 301 AACGAGCTCCACCTCTGCTGCTAGCAGCCGAGATGGGGAAATTTACATCAGAGGCT 360

QY 135 GlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGluLysValAlaValTyrLeuLys 154

DB 361 GGGCTTACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 155 LeuAspLeuLeuValAlaGlnValAlaLeuAlaLeuArgCysLeuGlnGluPheSerAlaThr 174

DB 421 CTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 175 AlaAlaSerSerProGlyProGlnLeuArgLeuGlnValSerGlyLeuLeuProLeu 194

DB 481 GCAGCAAGCTCTCTCTGGGCGCCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 195 ArgProGlySerSerLeuArgIleAlaThrLeuProTrpAlaHisLeuLysAlaAlaPro 214

DB 541 CGGCGAGGGTCTTCCCTTGGAGCGGACCTCCCTGGGCTCATTTAAGGCTGCCCC 600

QY 215 PheLeuThrTyrPheGlyLeuPheGlnValHis 225

DB 601 TTCTCAACCTACTTGGAGCTTTCAGAGTTGAC 633

RESULT 3

AAV47613

ID AAV47613 standard; cDNA; 1236 BP.

XX

AC AAV47613;

XX

DT 27-OCT-1998 (first entry)

XX

DE TNF related endothelium proliferative agent gene.

XX

KW ss; TNF: endothelium proliferative agent; TREPA: wound healing; cancer;

XX tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

XX

OS Homo sapiens.

XX

FX Key

FT CDS

FT

FT

FT

XX

PN WC9835061-A2.

XX

PD 13-AUG-1998.

XX

PF 12-FEB-1998; 98WO-US02859.

XX

PR 10-FEB-1998; 98US-0021706.

PR 12-FEB-1997; 97US-0798692.

XX

PA (ABBO ) ABBOTT LAB.

XX

PI Wiley SR;

XX

DR WPI: 1998-447255/38.

DR P-PSDB; AAM29745.

XX

PT Detecting nucleic acid encoding TREPA - useful for diagnosis and

PT treatment of autoimmune disease, tumours and inflammation

XX

PS Claim 11; Page 123-4; 142pp; English.

XX

CC The TNF-related endothelium proliferative agent (TREPA), or its

CC activators or agonists, are used to treat a deficit of TREPA, e.g. to

CC promote wound healing or tissue grafting, by promoting vascularisation,

CC also to induce apoptosis for treating cancer and eliminating autoreactive

CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.

CC TREPA peptides can also be used to target cytotoxic agents or for

CC affinity isolation of the corresponding receptor, the nucleic acid for

CC which can be used to transform tumour cells to render them more

CC responsive to TREPA and to screen for TREPA mimics.

CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat

CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting

CC vascularisation), inflammation or a wide range of autoimmune conditions,

CC conditions involving abnormal stimulation of epithelial cells (e.g.

CC atherosclerosis), for birth control (inhibiting ovulation and placental

CC formation) or other angiogenic conditions (e.g. ulcers).

XX

SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Alignment Scores:

Pred. No.:	7.98e-84	Length:	1236
Score:	1020.00 <td>Matches:</td> <td>199 </td>	Matches:	199
Percent Similarity:	92.86% <td>Conservative:</td> <td>9 </td>	Conservative:	9
Best Local Similarity:	88.84% <td>Mismatches:</td> <td>16 </td>	Mismatches:	16

```

Query Match:      87.78%      Indels:      0
DB:              19          Gaps:         0
US-09-245-198a-2 (1-225) x AAV47613 (1-1236)

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 135

OY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 AGCGGGGACATGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 195

OY 42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 CAGGACCCGCGGACATGATCCCGACAGAGAAAGACAGATCTCGGCTTTCCTG 255

OY 62 GluGlnLeuValArgProArgArgSerAlaProGlyGlyArgGlyAlaArgProArgArg 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 AACCGACTAGTTCGCTGGCTGGAGAGATGCACTAAAGCGCGGAAACACGCGCTGCAAGA 315

OY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 GCGATGCGAGCCCATTTATGAAGTTTCATCCAGACTGGAGAGGAGCGAGCGAGCGGT 375

OY 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 376 GTGAGCGGAGACAGTGAAGTGGCTGGGAGAGAGCAAGATCAACAGCTCCAGCCCTGGCGC 435

OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 436 TACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTGCTACTACTGT 495

OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 496 CAGTGCACCTTGTATGAGGGAGGAGGCTGTCTACTGACAGCTGACCTTGGCTGGAGTGT 555

OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAsnSerProGlyPro 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 556 GTGCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 615

OY 182 GlnLeuArgLeuGlnGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 616 CAGCTCGGCTGCTGGAGGCTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCGG 675

OY 202 IleArgThrLeuProTrpAlaHisIleLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 676 ATCCGACACCTCCCTCCGCGGCCATCTCAAGCGTCCGCCCTTCTCCTCACTTCCGAGCTC 735

OY 222 PheGlnValHis 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 736 TTCAGGTTTCAC 747

RESULT 4
AAD04350 standard; cDNA: 1236 BP.
ID AAD04350
AC AAD04350;
XX
XX
XX 04-JUL-2001 (first entry)
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnerability; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..750
XX FT
XX /product= "Human TREPA (TNF related endothelium

```

```

FT FT      proliferative agent)"
PN PN      US6207642-B1.
XX XX      27-MAR-2001.
XX XX      26-JUN-1998; 98US-0105343.
XX XX      12-FEB-1997; 97US-0798692.
XX XX      10-FEB-1998; 98US-0021706.
XX XX      (ABBO ) ABBOTT LAB.
XX XX      WILEY SR.
XX XX      WPI: 2001-280760/29.
XX XX      F-PSDB; ABE00891.
XX XX      Inducing angiogenesis in mammal at desired sites for promoting wound
XX XX      healing, by administering soluble fragment of extracellular domain of
XX XX      tumor necrosis factor related endothelium proliferative agent protein
XX XX      -
XX XX      Example 2; Column 73-74; 53pp; English.
XX XX
XX XX      The present invention relates to extracellular signal molecules,
XX XX      particularly members of tumour necrosis factor (TNF) family molecules
XX XX      designated as TREPA (TNF related endothelium proliferative agent).
XX XX      Soluble biologically active TREPA are used to treat TREPA-associated
XX XX      diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX XX      in human for promoting wound healing and for vascularising grafted
XX XX      tissue for successful grafting and to promote tissue grafts.
XX XX      The present sequence is a cDNA clone ID #690050 encoding human TREPA.
XX XX
XX XX      Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:
XX XX
XX XX      Alignment Scores:
XX XX      Pred. No.: 7.98e-84      Length: 1236
XX XX      Score: 1020.00      Matches: 199
XX XX      Percent Similarity: 92.86%      Conservative: 9
XX XX      Best Local Similarity: 88.84%      Mismatches: 16
XX XX      Query Match: 87.78%      Indels: 0
XX XX      DB: 22      Gaps: 0

US-09-245-198a-2 (1-225) x AAD04350 (1-1236)

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 135

OY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 AGCGGGGACATGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 195

OY 42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 CAGGACCCGCGGACATGATCCCGACAGAGAAAGACAGATCTCGGCTTTCCTG 255

OY 62 GluGlnLeuValArgProArgArgSerAlaProGlyGlyArgGlyAlaArgProArgArg 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 AACCGACTAGTTCGCTGGCTGGAGAGATGCACTAAAGCGCGGAAACACGCGCTGCAAGA 315

OY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 GCGATGCGAGCCCATTTATGAAGTTTCATCCAGACTGGAGAGGAGCGAGCGAGCGGT 375

OY 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 376 GTGAGCGGAGACAGTGAAGTGGCTGGGAGAGAGCAAGATCAACAGCTCCAGCCCTGGCGC 435

OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 436 TACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTGCTACTACTGT 495

```

QY 142 GlnValHisPheAspGluLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 496 CAGGTGACACTTGGATGAGGGGAAAGGCTGCTCTACCTGAACTGACTGCTGGTGGATGGT 555  
 QY 162 ValLeuAlaLeuArgCysLeuGluLuphSerAlaThrAlaLaserSerProGlyPro 181  
 DB 556 GTGCTGGCCCTGCTGCTGGATGAGAAATCTCAGCCACTGGCGGAGTTCCTCGGGGCC 615  
 QY 182 GlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 616 CAGCTCCGCTGCGAGGTGCTGGGCTGGCCCTGGCCGAGGGTCTCCCTGCGG 675  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
 DB 676 ATCCGACACCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCACTTCCGAGCTC 735  
 QY 222 PheGlnValHis 225  
 DB 736 TTCAGGTTTAC 747  
 RESULT 5  
 AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 AC AAA49717;  
 XX 25-SEP-2000 (first entry)  
 DE Human PRO207 cDNA clone DNA30879-1152.  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide 58..177  
 FT /\*tag= a  
 FT /\*tag= b  
 FT mat\_peptide 178..804  
 FT /\*tag= c  
 FT  
 PN WO200037638-A2.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99WO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 PA (GENE) GENENTECH INC.  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pitti RM, Wood WI;  
 DR MPI; 2000-442668/38.  
 DR P-PSDB; AAY95338.  
 XX Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or

PT PRO866 -  
 XX  
 PS Claim 20; Fig 3; 172pp; English.  
 XX  
 CC The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.4%). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,  
 CC PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see  
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.  
 CC  
 SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 8,95e-84 Length: 1353  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-245-198A-2 (1-225) x AAA49717 (1-1353)  
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
 DB 133 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192  
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluLuphSerGlnGluLeuThrAlaGluAspArg 41  
 DB 193 ACCCGGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252  
 QY 42 ArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAspValProPheLeu 61  
 DB 253 CAGGACCGCTGCGAACTGATCCAGAGAGAAAGGACAGAGATCTGCGCTTCTCG 312  
 QY 62 GlnGlnLeuValArgProArgArgSerAlaProArgArgArgArgArgArgArgArg 81  
 DB 313 AACGACATGTTGCGCTGCGAGAGTGCACCTAAAGCGGAGAAACACAGGGGCTGAGAA 372  
 QY 82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGlyValGlnAlaGly 101  
 DB 373 GCGATCGCAGCCCATTAATCAATTCATCCAGACCTGGACAGAGCGAGCGCAGCGAGT 432  
 QY 102 ValAspGlyThrValSerGlyTyrPglGluLuphLysIleAsnSerSerProLeuArg 121  
 DB 433 GTGAGCGGAGACAGTAGTGTGGGAGAGAGCAATCAACAGCTCCAGCCCTCGGC 492  
 QY 122 TyrAspArgGlnIleGlyGluLuphThrValIleArgAlaGlyLeuTyrTyrLeuArg 141  
 DB 493 TGCACACCGGACAGATCGGGAGGTTATAGCACCGGGCTGGCTTACACCTGTAAGT 552  
 QY 142 GlnValHisPheAspGluLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 553 CAGGTGACACTTGGATGAGGGGAAAGGCTGCTCTACCTGAACTGACTGCTGGTGGATGGT 612  
 QY 162 ValLeuAlaLeuArgCysLeuGluLuphSerAlaThrAlaLaserSerProGlyPro 181  
 DB 613 GTGCTGGCCCTGCTGCTGGATGAGAAATCTCAGCCACTGGCGGAGTTCCTCGGGGCC 672  
 QY 182 GlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 673 CAGCTCCGCTGCGAGGTGCTGGGCTGGCCCTGGCCGAGGGTCTCCCTGCGG 732  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221



Db	733	ATCCGACACCTCCCTGGGCCATCTCAGAGCGTCCCTCTCCTACCTTGGAATC	792
Oy	222	PheGlnValHis 225	
Db	793	TTCCAGGTTTCAAC 804	
RESULT 6			
ABK40255			
ID	ABK40255	standard; cdna; 1353 BP.	
XX			
AC	ABK40255;		
XX			
DT	15-JUL-2002	(first entry)	
XX			
DE	cdna encoding human PRO207 polypeptide.		
XX			
KW	Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;		
KW	leukemia; neuronal disorder; stromal disorder; blastocoele disorder;		
KW	inflammatory disorder; immune disorder; angiogenic disorder;		
XX	gene therapy; cytostatic; neuroprotective; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153486-A1.		
PD			
XX	26-JUL-2001.		
XX			
PF	11-FEB-2000; 2000WO-US03565.		
XX			
PR	08-MAR-1999; 99WO-US05028.		
PR	11-MAR-1999; 99US-123972P.		
PR	11-MAY-1999; 99US-133459P.		
PR	02-JUN-1999; 99WO-US12252.		
PR	22-JUN-1999; 99US-140650P.		
PR	22-JUN-1999; 99US-140653P.		
PR	20-JUL-1999; 99US-144758P.		
PR	26-JUL-1999; 99US-145698P.		
PR	28-JUL-1999; 99US-146222P.		
PR	17-AUG-1999; 99US-149385P.		
PR	31-AUG-1999; 99US-151689P.		
PR	01-SEP-1999; 99WO-US20111.		
PR	15-SEP-1999; 99WO-US21090.		
PR	30-NOV-1999; 99WO-US28313.		
PR	01-DEC-1999; 99WO-US28301.		
PR	01-DEC-1999; 99WO-US28634.		
PR	05-JAN-2000; 2000WO-US00219.		
XX			
PA	(GETH ) GENENTECH INC.		
XX			
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;		
PI	Morsters SA, Pan J, Pittt RM, Roy MA, Smith V, Stone DM;		
PI	Ketanabe CK, Wood WJ;		
XX			
DR	WPI; 2002-205567/26.		
DR	P-PSDB; AA086129.		
PT	Thirty five nucleic acids encoding PRO polypeptides, useful for		
PT	treating benign or malignant tumours, leukemias and lymphoid		
PT	malignancies, inflammatory, angiogenic and immunologic disorders -		
XX			
XX	Claim 50; Fig 3; 302pp; English.		
XX			
CC	The present invention relates to the isolation of novel human PRO		
CC	polypeptides and the polynucleotide sequences encoding them. The		
CC	PRO polypeptides, agonists, antagonists or anti-PRO antibodies are		
CC	useful for treating benign or malignant tumours (e.g. renal, kidney,		
CC	bladder, breast, etc), leukemias and lymphoid malignancies, other		
CC	disorders such as neuronal, glial, astrocytal, hypothalamic, glandular		
CC	macrophagal, stromal and blastocoele disorders, inflammatory, immune		
CC	and angiogenic disorders. The polynucleotide sequences are also		
CC	useful in gene therapy. ABK40254 ABK40288 encode for the human PRO		
CC	polypeptides of the invention.		

[illegible]



KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;  
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
 KW tissue regeneration; wound healing; burn; haematopoiesis;  
 KW myeloid cell deficiency; lymphoid cell deficiency.

OS Homo sapiens.

PN WO200177288-A2.

PD 18-OCT-2001.

PE 29-MAR-2001; 2001WO-US10224.

PR 06-APR-2000; 2000US-195582P.

PA (GEMV ) GENETICS INST INC.

PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;  
 PI Gulikova K, Graham JR;

DR WPI: 2002-179321/23.

PT Five hundred and ninety two polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders

PS Claim 1; Page 82; 372pp; English.

XX The invention relates to 592 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins. The polynucleotides can be used as probes for the  
 CC identification and isolation of full length cDNA and genomic DNA. The  
 CC polynucleotides and proteins can also be used as nutritional supplements.  
 CC The proteins are useful in the treatment of various immune deficiencies  
 CC and disorders such as viral infections, bacterial infections, fungal  
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
 CC and conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment of  
 CC burns, incisions and ulcers. The proteins are also useful for regulating  
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention.  
 XX

SQ Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;

Alignment Scores:

Pred. No.: 9.04e-84 Length: 1364  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 24 Gaps: 0

US-09-245-198a-2 (1-225) x ABK34881 (1-1364)

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGly 21  
 DB 161 CTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGG 220  
 OY 22 SerTTPAlaThrLeuSerLeuGlyLeuProSerGlnGlnGlnLeuThrAlaGlnAspArg 41  
 DB 221 ACCCGGCGATCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGG 280  
 OY 42 ArgGluProGlnLeuLeuAsnProGlnThrGlnGlnGlnSerGlnAspValValProPheLeu 61  
 DB 281 CAGGACCGCTCGAATCCAGACGAGAAAGCCAGAGATCCCTTCCTTCCTG 340

OY 62 GlnGlnLeuValArgProArgArgSerAlaProGlyGlyArgValAlaArgProArgArg 81  
 DB 341 AACCACTAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGG 400  
 OY 82 AlaIleAlaAlaIleTyrGlnValAlaIleProArgProGlyGlnAspGlyValAlaGln 101  
 DB 401 GCGATCGGACGCCATTATGAACTTCATCCAGACCTGGACAGAGAGGAGCGAGGAGGT 460  
 OY 102 ValAspGlyThrValSerGlyTyrProGlnThrIleValIleAsnSerSerProLeuArg 121  
 DB 461 GTGGACGGGACAGTAGTGGCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 520  
 OY 122 TyrAspArgGlnIleGlyGlnPheThrValIleArgAlaGlyLeuThrTyrLeuTyrGly 141  
 DB 521 TACAAACCGGACATCGGGAGGAGTTTATGATCACCAGGCTGTGGCTGTGGCTGTGGCT 580  
 OY 142 GlnValIleHisPheAspGlnGlyGlyValAlaValTyrLeuValLeuAspLeuValAlaGly 161  
 DB 581 CAGGTGCACTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 640  
 OY 162 ValIleAlaLeuArgCysLeuGlnGlnPheSerAlaThrAlaIleSerSerProGlyPro 181  
 DB 641 GTGCTGGCCCTGCGCTGTGGAGGAAATTCAGCCACTGCGGAGGAGTCCCTCGGGCCC 700  
 OY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 701 CAGCTCGGCTGTGGCAGGATGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGG 760  
 OY 202 IleArgThrLeuProTyrAlaIleIleValAlaIleProPheLeuThrTyrPheGlyLeu 221  
 DB 761 ATCCGACCCCTCCCTGGGCGCATCTCAAGGCTGCGCCCTTCTCACTTCCGAGATC 820  
 OY 222 PheGlnValHis 225  
 DB 821 TTCAGAGTTTAC 832  
 RESULT 8  
 AAV18600  
 ID AAV18600 standard; cDNA, 1373 BP.  
 XX  
 AC AAV18600;  
 XX  
 XX 21-JUL-1998 (first entry)  
 DE Homo sapiens tumour necrosis factor related ligand (TRELL) gene.  
 XX  
 XX TRELL; tumour necrosis factor related ligand; tnfr; treatment;  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..852 /\*tag= a  
 FT /note= "tumour necrosis factor related ligand"  
 FT  
 XX WO9805783-A1.  
 PN 12-FEB-1998.  
 PD 07-AUG-1997; 97WO-US13945.  
 PF 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 PA (BIOT ) BIOGEN INC.  
 PA (UYCE-) UNIV GENEVA FACULTY MEDICINE.  
 XX Browning JL, Chicheportliche Y;  
 XX WPI: 1998-145619/13.

DR P-PSDB; AAN47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
XX Claim 2: Pages 48-50; 69pp; English.

XX The sequence is that encoding human tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of prepare probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.

XX Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:

#### Alignment Scores:

Pred. No.:	9.12e-84	Length:	1373
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	19	Gaps:	0

US-09-245-198a-2 (1-225) x AAV18600 (1-1373)

```

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
DB 181 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 240
OY 22 SerTPAlaThrLeuSerAlaGlnGlnProSerGlnGlnGlnGlnGlnGlnGln 41
DB 241 AGCCGGCATGCTCTCCGCCAGAGACCTGCCAGAGAGAGAGAGAGAGAGAGAG 300
OY 42 ArgGluProGluLeuAsnProGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGln 61
DB 301 CAGGACCCGTCGAACTCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 62 GluGlnLeuValArgProArgSerAlaProGlySerGlyAlaArgProArgArg 81
DB 361 AACCGACAGTTCGGCTCGCAGAGTCCACCTAAAGCCCGAAMACAGGCGCTCGAGA 420
OY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 101
DB 421 GCGATCGGAGCCCATTAAGACTTCATCCAGACTGACAGAGAGAGAGAGAGAGAG 480
OY 102 ValAspGlyThrValSerGlyTTPolGluThrTyrHisIleAsnSerSerProLeuArg 121
DB 481 GTGACGGGACAGTACAGTGGCTGGAGAGAACCCAGAACTCAACAGCTCCAGCTCGGC 540
OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
DB 541 TACAACCGCCAGATCGGGAGATTATAGTACCCGGGCTGGGCTTACTACTACTACTGT 600
OY 142 GlnValHisPheAspGlyGlyValAlaValTyrLeuValAspLeuValValAspGly 161
DB 601 CAGTGCACCTTTGATGAGGGAGAGAGCTCTCTACTGAAAGCTGGAGCTGTGGAGTGT 660
OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
DB 181

```

```

DB 661 GTCTGCCCCCTGCGCTGCTGGAGAGATTCACGCCATGCGGCGAGTTCCCTCGGGCCC 720
OY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
DB 721 CAGCTCCGCTCTGGCAGAGTGTCTGGGCTGTGGCCCTCGGCGAGAGGTCCTCCCTGGCG 780
OY 202 IleArgThrLeuProThrAlaHisIleuLysAlaIleProPheLeuThrTyrPheGlyLeu 221
DB 781 ATCCGACCCCTCCCTGGGCGCCATCTCAAGCTGCCCCCTCTCTCACTTCCGAGCTC 840
OY 222 PheGlnValHis 225
DB 841 TTCACAGTTTAC 852

```

#### RESULT 9

AAK56000  
ID AAK56000 standard; DNA: 1421 BP.

XX AAK56000;

XX 15-JUL-1999 (first entry)

XX Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.

XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;

XX NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 92..841

XX FT /\*tag= a

XX PN /product= "Apo-3 ligand"

XX MO9919490-A1.

XX 22-APR-1999.

XX 09-OCT-1998; 98WO-US21407.

XX 17-DEC-1997; 97US-0069862.

XX 10-OCT-1997; 97US-0062037.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Marsters SA, Pitti R;

XX WPT: 1999-287982/24.

XX P-PSDB; AAY09369.

XX New human Apo-3 ligand (a tumour necrosis factor) homologue

XX Claim 18; Fig 1; 74pp; English.

XX The present sequence encodes a human tumour necrosis factor (TNF) and

XX lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has

XX cytosolic activity. Apo-3 ligand can be used to induce apoptosis in

XX mammalian cancer cells; to induce NF-kappaB-dependent transcription and

XX to induce JNK/SAPK-dependent responses in mammalian cells.

#### Alignment Scores:

Pred. No.:	9.53e-84	Length:	1421
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-2 (1-225) x AAK56000 (1-1421)



```

QY 59 ProPheLeuGluGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg 78
DB 370 CCTTCTCTGACACCTAGTTCGGCTCGCAGAAAGTCAACCTAAAGCCGAGAAACACGG 429
QY 79 ProArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 98
DB 430 GCTCGAGAGCCGATCGCACCCCTTATGAAATTCAATCCAGACTGAGACGAGCAGCG 489
QY 99 GlnAlaGlyValAspGlyThrValSerGlyTyrGlyGluThrLysIleAsnSerSer 118
DB 490 CAGGAGGTGTGGACGGAGCAGATGGCTGGGAGGAAACCAAGATCAACAGCTCCAGC 549
QY 119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgLysLeuTyrTyr 138
DB 550 CCTGTGGCTACAAACCCGACGATCGGGAGATTATAGTCAACCCGGGCTGGGCTCTAC 609
QY 139 LeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeu 158
DB 610 CTGTACTGTCAAGTGCATTTATGATGAGGGAAGGCTGTCTACTGAACTGAGACTTGTG 669
QY 159 ValAsnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleAsnSer 178
DB 670 GTGATGTGTGTCTGGCCCTCGCGCTCGGAGGAATTCTCAGCCACTGCGGCACTTCC 729
QY 179 ProGlyProGlnLeuArgLysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
DB 730 CTCGGGCCCCACGCTCCGCTCGCCAGAGGTCTGTGGGCTGTGGCCCTGGCGGCAAGGCTCC 789
QY 199 SerLeuArgGlyLeuArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyr 218
DB 790 TCCCTGGGAGATCCGACACCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCTCACTAC 849
QY 219 PheGlyLeuPheGlnValHis 225
DB 850 TTCGGACTCTTCACAGTTCAC 870
RESULT 11
AAK23424
ID AAK23424 standard; DNA; 1030 BP.
XX
AC AAK23424;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..627
FT /tag= a
FT /product= "TNRL3"
XX
PD MO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98MO-US18393.
XX
PR 05-SEP-1997; 97US-0924634.
XX
PA (UNIT) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI; 1999-205191/17.
DR P-PSDB; AAW93590.

```

```

XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Example VII: Fig 13A; 156pp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other.
XX
Alignment Scores:
Pred. No.: 1.26e-77 Length: 1030
Score: 951.00 Matches: 185
Percent Similarity: 92.79% Conservative: 8
Best Local Similarity: 88.94% Mismatches: 15
Query Match: 81.84% Indels: 0
DB: Gaps: 0
US-09-245-198a-2 (1-225) x AAK23424 (1-1030)
QY 18 ValSerLeuGlySerTyrPalaThrLeuSerAlaGlnGluProSerGlnGluLeuThr 37
DB 1 CTCAGTTGGGGAGCCGGGACATCGCTGCCCCCAAGGAGCTGCCAGAGACTGGTG 60
QY 38 AlaGluAspArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspVal 57
DB 61 GCAGAGGAGGACGAGACCCGTCGGAAGTCAATCCAGACAGAAAGCCAGAGATCT 120
QY 58 ValProPheLeuGluGlnLeuValArgProArgArgSerAlaProLysGlyAlaGlyAla 77
DB 121 GCGCTTCTCGAACCACTAGTTGCGCTCGCAGAAAGTGCACCTAAAGCCGGAACA 180
QY 78 ArgProArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGly 97
DB 181 CGGCTCGAAGAGCGATCGCAGCCATTAATGAACTTATCCAGACCTGAGACGAGGA 240
QY 98 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrGlyGluThrLysIleAsnSerSer 117
DB 241 GCGCAGGACAGGTGTGGAGGACAGTGAAGTGGGAGGAGAACCAACCAACAGCTCC 300
QY 118 SerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr 137
DB 301 AGCCCTCTGCTACTACCCGCAAGTCCGGGAGATTATAGTCAACCCGGGCTGGCTTAC 360
QY 138 TyrLeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeu 157
DB 361 TACTGTACTGTCAAGGTGCATTTGATGAGGGAAGGCTGTCTTACTGAAAGTGGACTTG 420
QY 158 LeuValAsnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleAsnSer 177
DB 421 CTGTGTGATGTGTGTCTGGGCTCGCTGGAGGAATCTCAACCACTGCGGCACT 480

```

```

OY 178 SerProGlyProGlnLeuArgLeuGlySerGlnValSerGlyLeuLeuProLeuArgProGly 197
DB 481 TCCCTCGGGGCCCCACGCTCCGCTTCGCCAGGTGTGTGGGCTGTGGCCCTGGGGCCAGGG 540
OY 198 SerSerLeuArgGlyLeuArgProGlnLeuProGlnValSerGlnValAlaProPheLeuThr 217
DB 541 TCCCTCCGCGGATCCGACCCCTCCGCGGCGCCATCCAGCTGCGCCCTTCTCTCACC 600
OY 218 TyrPheGlyLeuPheGlnValHis 225
DB 601 TACTTCGGAGCTCTCCAGGTTTAC 624

RESULT 12
ABK11680
ID ABK11680 standard; DNA: 2148 BP.
AC
XX ABK11680;
XX
XX 05-JUN-2002 (first entry)
XX
DE DNA encoding tumour necrosis factor variant 2 (TNFV2).
XX
KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;
KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;
KW haemorrhagic necrosis of tumour; cytotoxicity; TNFV2;
KW obesity-linked insulin resistance; gene; ds.
XX
OS Drosophila melanogaster.
XX
XX
FH Key Location/Qualifiers
FH sig_peptide 634..789
FT /*tag= a
FT CDS 634..1863
FT /*tag= b
FT /product= "TNFV1"
FT /note= "Tumour necrosis factor variant 1"
FT mat_peptide 790..1860
FT /*tag= c
FT /label= mature_TNFV1
FT misc_difference 634..1860
FT /*tag= d
FT /note= "Specifically claimed in claim 22"
XX
XX US2002012968-A1.
XX
XX 31-JAN-2002.
XX
XX 20-MAR-2001; 2001US-0813329.
XX
XX 21-MAR-2000; 2000US-190816P.
XX
XX (CARR/) CARROLL P M.
XX (CHEN/) CHEN J.
XX (RAMA/) RAMANATHAN C S.
XX (XIAO/) XIAO H.
XX (GUAN/) GUAN B.
XX (BOWE/) BOWEN M A.
XX
XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA.
XX WPI: 2002-195121/25.
XX P-PSDB: AAU77718.
XX
XX New Drosophila tumour necrosis factor molecule, useful in controlling
XX agriculturally important pests, e.g. comprises modifying the growth,
XX feeding or reproduction of crop-damaging insects or insects of farm
XX animals -
XX
XX Claim 2; Fig 3A-C; 119pp; English.
XX

```

```

CC The invention describes an isolated tumour necrosis factor polypeptide
CC (TNF). The polypeptide and polynucleotide are useful in controlling
CC agriculturally important pests, particularly by modifying the growth,
CC feeding and/or reproduction of crop-damaging insects or insects of farm
CC animals. The polypeptide and polynucleotide are useful for modulating
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals.
CC Thus the polypeptide and polynucleotide may be useful for treating,
CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal
CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal
CC dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or
CC sweat gland aberrations in animals (e.g. insects and potentially humans),
CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,
CC cytotoxicity and obesity-linked insulin resistance, all of which involve
CC TNF molecules. This sequence encodes the drosophila melanogaster tumour
CC necrosis factor variant 2 (TNFV2) protein, described in the invention.
XX
XX
SQ Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 other;
XX
XX
Alignment Scores:
Pred. No.: 0.962 Length: 2148
Score: 113.50 Matches: 50
Percent Similarity: 41.54% Conservative: 31
Best Local Similarity: 25.64% Mismatches: 78
Query Match: 9.77% Indels: 36
DB: Gaps: 9
XX
US-09-245-198a-2 (1-225) x ABK11680 (1-2148)
OY 41 ArgArgGluProProGlnLeuArgProGlnValSerGlnValPro-Phe 60
DB 1353 CAAGAGAGAGCCCTGCACCTTCACCA-----CCGTCCAGAAATCATCTCCGCCCA 1403
OY 60 eLeuGlnLeuValArgProArgSerAlaProGlySerGlyAlaArgProArg 80
DB 1404 TCGCCACCTCTAGTCCGAAAGCCAGATCC-----GAGACTCCAGCCCA-- 1449
OY 80 gArgAlaIleAlaHisTyrGlyValHisProArgProGlyGlnAspGlyAlaGln 100
DB 1450 -----GCAGCCCATTTCCACTTGCAGCAGCGCGCGCTCCACCAAGAAATG-- 1497
OY 100 aGlyValAspGlyThrValSerGlyTyrGlnThrValSerSerSerProle 120
DB 1498 -GGTACCATGAGATATGATATGATGAAATGATTAACGAGAAACTGTATTCAGGGA-- 1554
OY 120 uArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeu 140
DB 1555 -CACTTCAACGCGCGATGGCTGTGACGGTACCAATACAGGCTATATTACGTATA 1613
OY 140 rCysGlnValHisPheAspGlyGlyLysAlaValIleValLysLeuAspLeuValAs 160
DB 1614 CGCCAGATATGCTACACAACTCCGACGAC-----CAGCA 1649
OY 160 nGlyValLeuAla-----LeuArgCysLeuGlnGluPheSerAl 173
DB 1650 CGGATTTTATCGCTTTCAAGAGACACATCCATCTCGAGTGTCTGAAC----- 1698
OY 173 aThrAlaIleSerSerProGlyProGlnLeuArgLeuGlySerGlnValSerGlyLeuLeu 193
DB 1699 -ACGTGCCACCATGCGACATGCGACATGCGACATGCGACATGCGACATGCGACATGCG 1757
OY 193 oLeuArgProGlySerSerSerLeuArgIleArgThrLeu-----ProTyrAlaHisIste 210
DB 1758 CCTGAGACGAACGAGAGATCATCTGAAGACATTCACCAACGATCCGAAATGCACTTCT 1817
OY 210 uLysAlaIleAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224
DB 1818 GCGGGAGGGAACAACGCAAGCTACTTGGCATCTTCAAGGTG 1860

RESULT 13
ABK11679
ID ABK11679 standard; DNA: 1221 BP.
AC
XX ABK11679;
XX

```

XX 05-JUN-2002 (first entry)

DT DNA encoding tumour necrosis factor variant 1 (TNFV1).

DE

XX

XX Tumour necrosis factor: TNF; pesticide; crop-damaging insect;

KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;

KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;

KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;

KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;

KW haemorrhagic necrosis of tumour; cytotoxicity; TNFV1;

KW obesity-linked insulin resistance; gene; ds.

OS Drosophila melanogaster.

XX

XX Key Location/Qualifiers

FT sig\_peptide 1..156

FT /\*tag- a

FT CDS 1..1221

FT /\*tag- b

FT /\*product= "TNFV1"

FT /\*note= "Tumour necrosis factor variant 1"

FT mat\_peptide 157..1218

FT /\*tag- c

FT /\*label= mature\_TNFV1

FT misc\_difference 1..1218

FT /\*tag- d

FT /\*note= "Specifically claimed in claim 18"

FT

XX US2002012968-A1.

XX

XX 31-JAN-2002.

XX

XX 20-MAR-2001; 2001US-0813329.

XX

XX 21-MAR-2000; 2000US-190816P.

XX

XX (CARROL) CARROLL P M.

PA (CHEN) CHEN J.

PA (RAMA) RAMANATHAN C S.

PA (XIAO) XIAO H.

PA (GUAN) GUAN B.

PA (BOWEN) BOWEN M A.

XX

PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;

XX

DR MPL: 2002-195121/25.

XX

XX New Drosophila tumour necrosis factor molecule, useful in controlling

PT agriculturally important pests, e.g. comprises modifying the growth,

PT feeding or reproduction of crop-damaging insects or insects of farm

PT animals

XX

XX Claim 2; Fig 2A-B; 119p; English.

XX

XX The invention describes an isolated tumour necrosis factor polypeptide

CC (TNF). The polypeptide and polynucleotide are useful in controlling

CC agriculturally important pests, particularly by modifying the growth,

CC feeding and/or reproduction of crop-damaging insects or insects of farm

CC animals. The polypeptide and polynucleotide are useful for modulating

CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals.

CC Thus the polypeptide and polynucleotide may be useful for treating,

CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal

CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal

CC dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or

CC sweat gland aberrations in animals (e.g. insects and potentially humans),

CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,

CC cytotoxicity and obesity-linked insulin resistance, all of which involve

CC TNF molecules. This sequence encodes the drosophila melanogaster tumour

CC necrosis factor variant 1 (TNFV1) protein, described in the invention.

XX

XX Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 other;

Alignment Scores:

Pred. No.: 0.792 Length: 1221

Score: 111.00 Matches: 44

Percent Similarity: 39.81% Conservative: 38

Best Local Similarity: 21.36% Mismatch: 84

Query Match: 9.55% Indels: 40

DB: 24 Gaps: 8

US-09-245-198a-2 (1-225) x ABK11679 (1-1221)

QY 29 GlnGluProSerGlnGlnGlnLeuThrAlaLeuSparGArgGluProProGlnLeuAsn 48

DB 691 CAGGAAAGATCATCCAAAGAGGACCACTTCCAAAGAGATGATCCCGCATCGCCAC 750

QY 49 ProGlnThrGlnGlnSerGlnAspValAlaProPheLeuGlnGlnLeuValArgProArg 68

DB 751 CTCCTAGTCCGCAAGAGGATCTCTT----- 780

QY 69 ArgSerAlaProLysGlyArgLysAlaArgProArgAlaAlaAlaAlaHisTyrGlu 88

DB 781 -----TCAGCCAGATCCGAGACTCGAGGCCA-----GCAGCCCATTTCCAC 822

QY 89 ValHisProArgProGlnGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGly 108

DB 823 TTGACGACGACGCGCGCTCACCAAGAGATGTG---GGCTACCATGGAGATATGTACATA 879

QY 109 TrpGlnGlnThrLysIleAsnSerSerProLeuArgTyrAspArgGlnIleGlyGlu 128

DB 880 GGAATGATATACGAGAGAACTTTATCAGGA---CACTTCAACAGCGGATGGCTC 936

QY 129 PheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGly 148

DB 937 TTGACGGTGACCAATACAGGCTATATACCTATATACCCCGAGATATCTCAACACTCG 996

QY 149 LysAlaValTyrLeuLysLeuAspLeuValAsnGlyValLeuAla----- 164

DB 997 CACGAC-----CAGACGATTTATCTGCTTTCAAGAGAC 1032

QY 165 -----LeuArgCysLeuGlnGlnPheSerAlaThrAlaAlaSerSerProGlyPro 181

DB 1033 ACTCATTCCTCGAGCTGTTCAAC-----ACGGTCCCAACCAACATCCCAT 1080

QY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201

DB 1081 AAGGTGCACACCTGCCACAGAGGTGTGTATCCACTGGAAACGAAACGAGAGATCAT 1140

QY 202 IleArgThrLeu-----ProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyr 218

DB 1141 CTGAAGGACATTCACAAAGATCGCAATGCACTTTCGGGAGGAGAAACCAACGAGCTAC 1200

QY 219 PheGlyLeuPheGlnVal 224

DB 1201 TTTCGATCTTCAAGTGTG 1218

RESULT 14

ABL21473

ID ABL21473 standard; DNA: 978 BP.

XX

XX ABL21473:

AC

XX

XX 26-MAR-2002 (first entry)

XX

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.

DE

XX

XX Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

OS Drosophila melanogaster.

XX

XX WO200171042-A2.

XX

XX 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS  
 PS Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB1616175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIGO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.663 Length: 978  
 Score: 110.50 Matches: 49  
 Percent Similarity: 41.62% Conservative: 33  
 Best local Similarity: 24.87% Mismatches: 81  
 Query Match: 9.51% Indels: 34  
 DB: 23 Gaps: 9

US-09-245-198a-2 (1-225) x ABL21473 (1-978)

QY 41 ArgArgGluProGluLeuAsnProGlnThrGluGlnSerGlnAspValPro-PH 60  
 DB 450 CAAGAGAGAGCCCTGCACCATTCACCA-----CCGTGCAGAAATGATCCCGCA 500  
 QY 60 GluGluGlnLeuValaArgProArgArgSer-----AlaProLysGlyArgLysAlaAr 78  
 DB 501 TCGCCACCTCTAGTCCGCAAGGTGAATCTCTTCTTACGCGCAGATCCGAGACTCGAG 560  
 QY 78 gProArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAl 98  
 DB 561 GCCA-----GCAGCCCATTTCCACTTGAGCAGCAGCGCGCTCACCAGAGAG 608  
 QY 98 aglnAlaGlyValAspGlyThrValSerGlyTyrGlnGlnThrLysIleAsnSerSer 118  
 DB 609 TATG---GGTACCATGATGATGTACATAGGAATGATTAACAGAGAACTTTATCA 665  
 QY 118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrT 138  
 DB 666 GGA---CACCTTCAAAAGCGCGATGGCTGTGAGCGTGAACCAATACAGGCTTATATTA 722  
 QY 138 rLeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeu 158  
 DB 723 CGTATACCCCAAGATATGCTTACACAACTCCGACAGC----- 759  
 QY 158 uValAsnGlyValLeuAla-----LeuArgCysLeuGluGluP 171  
 DB 760 -CAGAACGATTTATGCTTTCAGAGAGACATCCCATTTCCGAGGTGTTAAAC----- 813  
 QY 171 eSerAlaAlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnValSerGlyLe 191  
 DB 814 -----ACGGTGCACCAACATGCGCAATAGGTGACACACTGCGACAGAGTGTCT 866

QY 191 uLeuProLeuArgProGlySerSerLeuArgIleArgThrLeu-----ProTyrAl 208  
 DB 867 GATCCACCTGGACAAACAGAGAGATTCATCTGAAGACATTCACACGATCGCAATGC 926  
 QY 208 aHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224  
 DB 927 AGTTCTCGCGGAGGAGAAACACCAAGCTACTGTCATCTTCAAGGTG 975  
 RESULT 15  
 ID AAV41377 standard; cDNA: 1630 BP.  
 XX AAV41377;  
 AC  
 XX 08-OCT-1998 (first entry)  
 DE NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.  
 XX  
 KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;  
 KW immune response; inflammatory response; toxic shock; sepsis;  
 KW RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.  
 OS Mus musculus.  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 3..887  
 FT /tag= a  
 FT /product= "murine RANKL (ligand for RANK)"  
 FT  
 XX WO9828426-A2.  
 XX  
 XX 02-JUL-1998.  
 PD  
 XX 22-DEC-1997; 97WO-US23775.  
 PF  
 XX 14-OCT-1997; 97US-0064671.  
 PR 23-DEC-1996; 96US-0059978.  
 PR 07-MAR-1997; 97US-0813509.  
 XX  
 PA (IMMEX ) IMMEX CORP.  
 XX  
 PI Anderson DM, Gallbert LJ, Maraskovsky E;  
 XX WPI; 1998-377657/32.  
 DR P-PSDB: AAM69956.  
 DB  
 XX  
 PT New isolated ligand for receptor activator of NF-kappa B - used to  
 PT develop products for augmenting an immune response for inhibiting an  
 PT inflammatory response and for protection of cells  
 PT  
 PS Claim 25; Pages 55-57; 80pp; English.  
 PS  
 XX This cDNA encodes a murine RANKL, a ligand for the RANK (receptor  
 CC activator of necrosis factor-kappa B (NF-kB)) polypeptide. RANK is a  
 CC member of the tumour necrosis factor (TNF) family. A soluble RANK  
 CC may be used for inhibiting activation of NF-kB, by contacting a cell  
 CC expressing membrane-associated RANK with a soluble RANK which binds to  
 CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be  
 CC used to induce maturation of dendritic cells and enhance their  
 CC allo-stimulatory capacity, thereby augmenting an immune response. The  
 CC soluble RANK polypeptide composition may also be used for regulating an  
 CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists  
 CC may be useful in ameliorating negative effects of an inflammatory  
 CC response that result from triggering of RANK, e.g. in treating toxic  
 CC shock or sepsis, graft-versus-host reactions, or acute inflammatory  
 CC reactions. They can also be used in adjunct therapy for disease  
 CC characterised by neoplastic cells that express RANK. RANKL polypeptides  
 CC can also be used to identify inhibitors of RANK and thus inhibitors of  
 CC an inflammatory response, and also for protecting RANK-expressing cells  
 CC from the negative effects of chemotherapy or the presence of high levels  
 CC of TNF-alpha. The products can also be used for detection and drug  
 CC screening.

50 Sequence 1630 BP: 436 A: 355 C: 379 G: 460 T: 0 other:

## Alignment Scores:

Pred. No.:	2.95	Length:	1630
Score:	106.50	Matches:	67
Percent Similarity:	38.048	Conservative:	38
Best Local Similarity:	24.288	Mismatches:	106
Query Match:	9.178	Indels:	65
DB:	19	Gaps:	13

US-09-245-198a-2 (1-225) x AAV41377 (1-1630)

```

QY 3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGlySer 22
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 78 TTCATGTCCTTCCTGCCCCCTGCGGCTGGAGCTGGGCCAGCTGCTGCGACATCGCTCG 137
QY 23 TrpAlaThrLeuSerAlaGln---GluProSerGlnGluLeuThrAlaGlnSparg 41
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 138 TTCCTGTACTTCGAGCGAGATGATGATCTACAGA-----ATATCAGAAAGACAGC 188
QY 42 -----ArgGluProProGluLeuAsnProGlnThr 51
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 189 ACTCACTGCTTTATAGAACTCTGAGACTCCATGAAACGAGATTGCGAGACTCGACT 248
QY 52 GluGluSerGlnAspValValPro----- 59
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 249 CTGGAGAGTGAAGACACACTACCTGCTGCTGCTGCGAGAGGAAACACCTTCAGGGG 308
QY 60 -----PheLeuGluGlnLeuValArgProArgArg-----SerAlaProLys 73
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 309 GCCGTGAGAGAACTGCAACACATGTGGGGCCACAGCGCTTCTCAGAGACTCCAGCT 368
QY 74 -----GlyArgLysAlaArgProArgArgAlaIleAla 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 369 ATGATGAGAGCTCATGTTGATGATGCGCCAGCGAGCGCAAGCCCTGAGCCGCAATTT 428
QY 85 AlaHisTyrgluValHisProArgProGlyGlnAspGlyValAlaGlyValAspGly 104
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 429 GCACACCTCACCACATCAATGCTGCCAGCATCCCATGGGTCCCAT-----AAAGTC 479
QY 105 ThrValSerGlyTyrPgluGlu-----ThrLysIleAsnSerSerProLeu 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 480 ACTCTGCTCTTGTGTACACGATCGAGCGAGCTGGGCCAAGATCTTAACATGACG----- 533
QY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 534 -----TTAAGCAACGAAACTAAGGTTAAACCAAGATGGCTTCTATTACCTGTAC 584
QY 141 -----CysGlnValHisPheAspGluGlyLysAlaVal-----TyrLeuLys 154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 585 GCCAACAATTCCTTCGCGCATCATGAACATCGGAGAGCGTACCTACAGACTATCTTCAG 644
QY 155 LeuAspLeuLeuVal-----AsnGlyValLeuAlaLeuArgCysLeuGluGlu 170
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 645 CTGATGCTGTATGCTGTTAAACACAGCATCAAAATCCCAAGTTCATTAACCTGATGAA 704
QY 171 PheSerAlaThrAlaHisSerSerProGlyProGlnLeuArgLeu-----CysGlnVal 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 705 GGAGGGAGCAGAAACCTGCTGCGCAATTCGAATTCACCTTTATTCATTAATGTT 764
QY 189 SerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTyrAla 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 765 GGGGGATTTTCAAGCTCCGAGCTGGTGAGAAATTTAGACATTCAGTGTCCAAACCTTCC 824
QY 209 HisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 825 CTGCTGATCCGATCAAGATGCGAGCTACTTTGGGCTTTCANAAGTT 872

```

Search completed: May 8, 2003, 03:42:39  
Job time : 171.65 secs



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:12 ; Search time 1765.08 seconds

(without alignments)  
3709.824 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162  
Sequence: 1 VSLGLALACGLLLVWVSL.....PWAHLKAAFLYFGILFQVH 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_plus\_p2n.model -DEV=tlp  
-Q=/cgn2\_l/uspt0.spool/us09245198/runat.06052003\_170239.9615/app-query.fasta.1.846  
-DB=GenEmbl -QFMT=fastap -SUFFIX=irge -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdl -LIST=45  
-DOCCALIGN=200 -THR.SCORE=pt -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us09245198.qcgn.1.1.3071 -runat.06052003\_170239.9615 -NCPV=6 -ICPV=3  
-NO\_XLPPX -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pin:\*  
35: em.htg\_rod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	10	AF030100	AF030100 Mus muscu
2	1020	87.8	1236	6	AR140407	AR140407 Sequence
3	1020	87.8	1306	9	AF030099	AF030099 Homo sapi
4	1020	87.8	1353	6	AX201324	AX201324 Sequence
5	1020	87.8	1368	9	AF055872	AF055872 Homo sapi
6	951	81.8	898	6	AX180714	AX180714 Sequence
7	945.5	81.4	1651	9	BC019047	BC019047 Homo sapi
8	602.5	51.9	203083	2	AC069459	AC069459 Mus muscu
9	602.5	51.9	234182	10	AF603707	AF603707 Mouse DNA
10	549.5	47.3	177703	2	AC016876	AC016876 Homo sapi
11	549.5	47.3	215795	2	AC127470	AC127470 Pan trogl
12	545	46.9	177555	2	AC130192	AC130192 Sus scrofa
13	544.5	46.9	161428	2	AC126925	AC126925 Canis fam
14	536.5	46.2	138792	2	AC119115	AC119115 Rattus no
15	503	43.3	153553	2	AC126921	AC126921 Bos tauru
16	425.5	36.6	184006	2	AC098923	AC098923 Rattus no
17	207	17.8	203281	2	AC126237	AC126237 Canis fam
18	113.5	9.8	1656	3	AB073865	AB073865 Drosophila
19	110.5	9.5	1221	3	AF511923	AF511923 Drosophila
20	110.5	9.5	2159	3	AF521176	AF521176 Gallus ga
21	109.5	9.4	892	5	GGA243435	GA243435 Human DNA
22	109.5	9.4	108967	9	AL353138	AL353138 Streptomy
23	108.5	9.3	42210	1	SC1C2	SC1C2 Pseudomon
24	107.5	9.3	13243	1	AE004602	AE004602 Homo sapi
25	107	9.2	178262	2	AC061974	AC061974 Homo sapi
26	107	9.2	179383	2	AC060789	AC060789 Homo sapi
27	106.5	9.2	1630	6	AR156433	AR156433 Sequence
28	106.5	9.2	1630	6	AR164147	AR164147 Sequence
29	106.5	9.2	1630	6	AX147987	AX147987 Sequence
30	106.5	9.2	1694	6	AX451897	AX451897 Sequence
31	106.5	9.2	2225	10	AF019048	AF019048 Mus muscu
32	106.5	9.2	4034	6	AX200995	AX200995 Sequence
33	106.5	9.2	4034	6	AX267730	AX267730 Sequence
34	106.5	9.2	4894	6	AX200993	AX200993 Sequence
35	106.5	9.2	4894	6	AX267728	AX267728 Sequence
36	106.5	9.2	6976	6	AX200996	AX200996 Sequence
37	106.5	9.2	6976	6	AX267731	AX267731 Sequence
38	106.5	9.2	157988	2	AC096533	AC096533 Homo sapi
39	106	9.1	177485	2	AC090975	AC090975 Mus muscu
40	106	9.1	207585	2	AC073779	AC073779 Mus muscu
41	106	9.1	222037	2	AC073801	AC073801 Mus muscu
42	105.5	9.1	4412	9	AY070219	AY070219 Homo sapi
43	105.5	9.1	6305	9	H0MLYTOXBB	L10106 Homo sapien
44	105.5	9.1	24526	6	AX472617	AX472617 Sequence
45	105.5	9.1	62370	2	AL451008	AL451008 Homo sapi

RESULT 1

# ALIGNMENTS

AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS AF030100 Mus musculus TWEAK mRNA, partial cds.  
 DEFINITION AF030100  
 ACCESSION AF030100  
 VERSION AF030100.1 GI:2707220  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1168)  
 Chicheportriche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H., Hession, C., Garcia, I. and Browning, J.L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)  
 MEDLINE 98070415  
 PUBMED 9405449  
 REFERENCE 2 (bases 1 to 1168)  
 Chicheportriche, Y., Bixler, S., Tizard, R. and Browning, J.  
 Direct Submission  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA  
 JOURNAL  
 FEATURES  
 source Location/Qualifiers  
 1..1168  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_type="peritoneal macrophages"  
 <1..679  
 /note="ligand in the TNF family; secreted protein"  
 /codon\_start=2  
 /product="TWEAK"  
 /protein\_id="AAC3517.1"  
 /db\_xref="GI:2707221"  
 /translation="VLSGLALACGLLVVYSLGSNATLSAQBPSOBELEAEDREP  
 PEINPOTESODVPELEOLVPRPSAKGRARLAAHYEHPRPGDGAQAV  
 DGTYSAGMETKINSSPLRYRDRIGETVIRAGLYLYLCQVHFDEGKAVYLLDLVY  
 GVALRCLLEFSATRASSPGPOLRICOVSGLLPRSSLRIRLTPMAHLAAAPLTY  
 FGLFQVH"  
 CDS  
 BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,47e-79 Length: 1168  
 Score: 1162.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-245-198a-2 (1-225) x AF030100 (1-1168)  
 QY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValValSerLeu 20  
 DB 2 GTGCTGAGCCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 61  
 QY 21 GlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlnGluLeuThrAlaGlnAsp 40  
 DB 62 GGGAGCTGGGCAACGCTGCTGCTCCCGAGAGCCTTCACAGAGAGCTGACAGACAGAGAC 121  
 QY 41 ArgArgGluProGluLeuAsnProGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGln 60  
 DB 122 CCCCCGGACCCCTGAACTGAATCCCGACAGAGAAAGCAGAGATGTGTACCTTTTC 181  
 QY 61 LeuGlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArg 80  
 DB 182 TTGGAAACAACAGTACCGCGCTCGAACAAGTCTCTAAAGCCGCAAGCCGCGCTCCG 241  
 QY 81 ArgAlaIleAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100  
 DB 242 CGAGCTATGTCAGACCATATGAGTTCACTCGGCCAGAGACAGATGAGACAGACAGA 301  
 QY 101 GlyValAspGlyThrValSerGlyTrpGlnGluThrLysIleAsnSerSerSerProLeu 120

|||||  
 DB 302 GGTGTGATGGAGGACAGTGTAGTGGTGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361  
 QY 121 ArgTyrAspArgGlnIleLeuGlnPheThrValIleArgAlaGlyLeuLeuTyrTyrLeuTyr 140  
 DB 362 CGCTACGACCGCAGATTGTGGGAATTTACAGTATCAGGGGTGGGTCTACTACCTGAC 421  
 QY 141 CysGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAsn 160  
 DB 422 TGTCAAGTCACTTTATATAGGAAAGCGTGTCTACCTCAACCTGACCTTGTGGTGAAC 481  
 QY 161 GlyValLeuAlaLeuArgCysLeuGlnGluLysSerAlaThrAlaAlaSerProGly 180  
 DB 482 GGTGTCTGGCCCTGGCGCTGCTGGAAATTTCTACGCCACAGACAGCAAGCTCTCTGGG 541  
 QY 181 ProGlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerLeu 200  
 DB 542 CCCCAGCTCCGTTTGTGGCAGAGTGTGGCGTGTGGCGCTGGCGGCGGCGGCTTCCCTT 601  
 QY 201 ArgLeuArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGly 220  
 DB 602 CGGATCGGACACCTCCCTCGGCTCATCTTAAGGCTGCCCTTCTTACCTACTTGA 661  
 QY 221 LeuPheGlnValHis 225  
 DB 662 CTCTTCAAGTTCAC 676  
 RESULT 2  
 AR140407  
 LOCUS AR140407 1236 bp DNA linear PAT 16-JUN-2001  
 DEFINITION Sequence 1 from patent US 6207642.  
 ACCESSION AR140407  
 VERSION AR140407.1 GI:14482903  
 KEYWORDS  
 ORGANISM Unknown.  
 SOURCE Unknown.  
 REFERENCE Unclassified.  
 AUTHORS 1 (bases 1 to 1236)  
 TITLE Member of the TNF family useful for treatment and diagnosis of disease  
 JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
 FEATURES  
 source Location/Qualifiers  
 1..1236  
 /organism="unknown"  
 BASE COUNT 225 a 416 c 358 g 237 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.38e-69 Length: 1236  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-245-198a-2 (1-225) x AR140407 (1-1236)  
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValValSerLeuGly 21  
 DB 76 CTGGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 135  
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlnGlnGlnGlnGlnGlnGlnGln 41  
 DB 136 AGCCGGGCGATCGTGTCCGCCAGAGAGCCTGCCAGAGAGAGCTGGTGTGACAGAGAGAC 195  
 QY 42 ArgGluProGluLeuAsnProGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGln 61  
 DB 196 CAGGACCCCTCGAAGTGAATCCCGACAGAGAAAGCAGAGATCTGCGCTTCCCTG 255  
 QY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArg 81  
 DB 256 AACCGACTGTTGGCTCGAAGAGTGCACCTAAAGCCGGAAGAACACAGGCTCGAAGA 315

OY	82	AlaIleAlaIaAhStyTgclValHisProRprProGlyInaSpGlyAlaGlalactyl	101
Db	316	GCGATCCGAGGCCATTAGAAGTTCAACAGACTGGACAGAGCGAGCGAGCAGCT	375
OY	102	ValAspGlyThrValSerGlyTYrPGLuglunThryHisIleaSenSerSerProLeuArg	121
Db	376	GTGGACGGGACAGTAGTGCGTGGGAGGAACAACAAATCAACAGCTCCAGCCCTCTGGGC	435
OY	122	TyrAspArgGlnIleGlyGluPhenThrValIleArgIlaGlyLeuTYrTYrLeuTYrCys	141
Db	436	TACAAACCCCGAATCGGGAGATTATATGTCACCCGGCGTGGGCTCTACTACTGTACTGT	495
OY	142	GlnAlaHisPheAspGluGlyLysAlaValTYrLeuLYsIleAspLeuValasnGly	161
Db	496	CAGTGCCACTTTGATAGAGGGAGAGGCTCTCTACCTGAAAGCTGGACTCTCGTGGATAGT	555
OY	162	ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIaIaSerSerProGlyPro	181
Db	556	GTGCTGGGCCCTGGCGTGGCTGGAGAAATTCACGACCATGGGGGAGATTCCTCGGGGCC	615
OY	182	GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg	201
Db	616	CAGCTCCCCTCTGCGACAGTGTCTGGGCTGTGGGCCCTGCGGCCAGGGCTCTCCCTGGCG	675
OY	202	IleArgThrLeuProTrpAlaHisIleuLYsAlaIaIaProPheLeuThrTYrPheGlyLeu	221
Db	676	ATCCGACACCTTCCTCGGGGCCCATCTCAAGCTGCCCCCTTCTCTACTACTTGGACTC	735
OY	222	PheGlnValHis	225
Db	736	TTCCAGGTTCAC	747
RESULT_3	AF030099		
LOCUS	AF030099	1306 bp	mRNA linear PRI 20-DEC-1997
DEFINITION	Homo sapiens TWEAK mRNA, complete cds.		
VERSION	AF030099		
KEYWORDS	AF030099.1 GI:2707218		
SOURCE			
ORGANISM	Homo sapiens.		
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1306) Chichportliche Y., Bourdon P.R., Xu H., Hsu Y.M., Scott H., Hesston C., Garcia I. and Browning J.L.		
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis		
JOURNAL MEDLINE	J. Biol. Chem. 272 (51), 32401-32410 (1997)		
PUBMED	98070415 9405449		
REFERENCE	2 (bases 1 to 1306) Bourdon P., Hesston C., Rizard R. and Browning J.		
AUTHORS	Direct Submission Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA		
JOURNAL			
FEATURES			
source	location/Qualifiers		
	1..1306 /organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
	/map="17pl3"		
	/tissue-type="tonsil"		
	18..767 /note="ligand in the TNF family; secreted protein; start codon not verified experimentally"		
	/codon_start=1		
	/product="TWEAK"		
	/protein_id="AAC51923.1"		
	/db_xref="GI:2707219"		
	/translation="MAARSRRRGRGRGEPTALVPLVLGLALACLILLAVSLGSRSLDADEPADEELVAEDQDSELNPTDESDAPFLNRLVRRSAPKGRKR		
CDS			

[illegible]

REFERENCE	TITLE	JOURNAL
1 (bases 1 to 1353)		
Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J., Pittl, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I.		
Compositions and methods for the treatment of tumour		
Patent: WO 0153486-A 3 26-JUL-2001;		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

Alignment Scores:	
Pred. No.:	1,03e-68
Score:	1020.00
Percent Similarity:	92.86%
Best Local Similarity:	80.84%
Query Match:	87.76%
DB:	6
	Gaps: 0
	Length: 135
	Matches: 199
	Conservative: 9
	Mismatches: 16
	Indels: 0
	Gaps: 0

US-09-245-198A-2 (1-225) x AX201324 (1-1353)

```

OY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 313 AACCGACTAGTTCGGCTCGCAGAACTGCACCTAAAGCGCGAAACACGGCTCCAGACA 372
OY 82 AAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
Db 373 GCGATCGCAGCCCATATGATGATTCATCCACAGCCTGGACAGACGACGACGACGAGGT 432
OY 102 ValAspArgTyrValSerLysTyrProLysGluTyrHisLysSerSerSerProLeuArg 121
Db 433 GTGGACGGGACAGATGAGTGTGGTGGGAGGAGCCAGATACAGCTCCAGCCCTCTGCC 492
OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrGly 141
Db 493 TACACCGCCAGATCGCGGAGATTATAGTCACCCGGGCTGGGCTCTACTACTGTCTGT 552
OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 553 CAGGTGCACTTGATGAGGGGAGGAGGCTGTACTGACCTGAGCTGACTGTGGTGGATGGT 612
OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
Db 613 GTGGTGGCCCTCGCTGCTGGAGAACTTCACGACCTGCGCGAGTTCCCTCGGGCCC 672
OY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CAGCTCCGGCTCTGCCAGGTGTCTGGGCTGTGGCCCTGGCGGAGGCTCTCCCTGCCG 732
OY 202 IleArgTyrLeuProTyrPheAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db 733 ATCCGACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTCTCTCACTTCTGCGAGTC 792
OY 222 PheGlnValHis 225
Db 793 TTCAGAGTTCAC 804

RESULT 6
LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
1 (bases 1 to 898)
REFERENCE
AUTHORS Wiley, S.R.
TITLE Weak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
IMMUNEX CORPORATION (US)

FEATURES
source location/Qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGRTSLIAFGILCLPMLQESGATSSDDMKQIEDKIEBILS
KIYHENEIARIKKIIGERTSSISGASISAOEBOEIVAEEDDSSEINPTEES
ODPAFTLRLVRRRSAPKGRKTRARRIAAHTVEHPRGDDAGAGYDGVSGWEEA
KINSSPLRNQIGEFITVRAGLYLVCQVHFDEGKAVYLKLDLIVGVIALRGLEE
FSATPAASSIGDRLRCQVSGILALRPGSSLRITLPMALNKAPLTYFGLEOVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

```

Alignment Scores:

Pred. No.:	1,15e-63	Length:	898
Score:	951.00	Matches:	184
Percent Similarity:	93.24%	Conservative:	9

```

Best Local Similarity: 88.89% Mismatches: 14
Query Match: 81.84% Indels: 0
DB: 6 Gaps: 0

US-09-245-198a-2 (1-225) x AX180714 (1-898)
OY 19 SerLeuGlySerTyrPheLeuSerAlaGlnGluProSerGlnGluLeuThrAla 38
Db 250 AGTTTGGGAGCCGGGCACTGCTGTCCGCCAGACCTTGGCCAGAGAGACTGTGGCA 309
OY 39 GluAspArgGluProProGluLeuAsnProGlnThrGluGlnSerGlnAspValVal 58
Db 310 GAGGACAGCAGAGACCCGTCGGAAGTATCCCAACAGAAAGAACAGAGATCCTGCG 369
OY 59 ProPheLeuGlnGluLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg 78
Db 370 CCTTTCCTAACCCAGACTTATGCGCTCCACAAAGCACCATAAAGCCGGAACACCG 429
OY 79 ProArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 98
Db 430 GCTCGAGAGCCGATCCGACCCATTTATGAACTTATCCAGACTGACAGAGGAGCGG 489
OY 99 GlnAlaGlyValAspGlyThrValSerGlyTyrProGluGluTyrHisLysSerSerSer 118
Db 490 CAGGACAGTGTGGACGCGACAGTGAAGTGTGGAGAGAACCCAGATCAACAGCTCCAGC 549
OY 119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr 138
Db 550 CCTTCGCTACACCGCCAGATCGGGAGTTATAGTACCCGGGCTGGCTCTACTAC 609
OY 139 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 158
Db 610 CTGTACTGTCAGAGTGCATTTGATGAGGAGGAGGCTGTACTGTAAGCTGCACTTGCTG 669
OY 159 ValAsnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSer 178
Db 670 GTGGATGATGTGCTGCTCGCTGGGCTCTGGAGAACTTCACAGCCCTCGGCGCAAGTTC 729
OY 179 ProGlyProGlnLeuArgCysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
Db 730 CTCGGGCCCCACCTCCGCTGCGCAGGTGTGGCTGTGGCCCTGGCGGAGGCTCC 789
OY 199 SerLeuArgIleArgTyrPheLeuProTyrPheAlaHisLeuLysAlaAlaProPheLeuThr 218
Db 790 TCCCTCGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTAC 849
OY 219 PheGlyLeuPheGlnValHis 225
Db 850 TTCGACTCTTCAGAGTTCAC 870

RESULT 7
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, Similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1651)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIN-MGC Project URL: http://mgc.ncl.nih.gov
CONTACT MGC help desk
Email: cgapbs-tr@mail.nih.gov

```











```

* 99494 99593: gap of 100 bp
* 99594 111049: contig of 11456 bp in length
* 11050 111149: gap of 100 bp
* 11150 125020: contig of 13871 bp in length
* 125021 125120: gap of 100 bp
* 125121 145109: contig of 19989 bp in length
* 145110 145209: gap of 100 bp
* 145210 169458: contig of 24249 bp in length
* 169459 169558: gap of 100 bp
* 169559 177703: contig of 8145 bp in length.
* Location/Qualifiers
  source          1. 177703
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="RP11-186B7"
BASE COUNT      42862 a 45967 c 46568 g 41254 t 1052 others
ORIGIN
Alignment Scores:
Pred. No.:      9,71e-31      Length:      177703
Score:          549.50        Matches:      117
Percent Similarity: 63.02%    Conservative: 4
Best Local Similarity: 60.94%  Mismatches:  5
Query Match:    47.29%        Indels:      66
DB:             2            Gaps:          1

US-09-245-198a-2 (1-225) x AC016876 (1-177703)
QY 100 A1aG1yAlaSpG1yThrValSerG1yTrpGluGluThrIleAsnSerSerPro 119
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 37636 TCAGGTGTGACGAGGACAGTGTGCTGGAGGAGCCAGCAATCAACAGCTCCAGCCCT 37637
QY 120 LeuArgTyrIleAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeu 139
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37676 CTGGGCTACACCCGCGAGTCGGGAGTTTATGTCACCCGCGGCTCTTACTACTCTG 37617
QY 140 TTTCTGys-----
      |||||
Db 37616 TACTGTGAGTAACCCCATCTGCTGCATGTTGTTACCCAGTAAAGAGTGGCCAGAGGT 37457
QY 141 -----
      |||||
Db 37456 TTGCGAGAGAGTGGGGGACACAGCTACAGGGCTGGAGAGTGTGGGCTTGGTGGG 37397
QY 141 -----
      |||||
Db 37396 ATGGGATGCTGCTGCTGAGGAATTTGAAATTTGAGGCGAGGCGAGGCGAGGCTTG 37337
QY 142 -----
      |||||
Db 37336 ACTGGGCTGTGTCGCCACGCCAGGTGCACTTTGATGAGGGGAAAGCTGTCTACCTGAA 37277
QY 154 s1euaSp1euleuValAsnGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThr 174
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37276 GCTTGAGATTGCTGTGCTGTGCTGGCCCTGCTGCTGCTGAGGAAATTCACGACAC 37217
QY 174 fAlaAlaSerSerProG1yPrG1nIleuArgLeuGlyGlnValSerG1yLeuLeuProLe 194
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37216 TGGGCGAGTTCCTCGGGGCGCCAGCTCCGCTCGCAGGATGTCGTCTGTTGGCCCT 37157
QY 194 uARgProG1ySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuIlysaAlaApr 214
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37156 GCGGCGAGGTCCTCCTCGCGGATCCGACACCTCCCTGGGCGCATCTCAAGGCTGCCCC 37097
QY 214 oPhLeuThrTyrPheGlyLeuPheGlnValHis 225
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37096 CTTCTCACCCTCTCGGACTCTCCAGGTTCCAC 37063
RESULT 11
AC127470 AC127470 215795 bp DNA linear HTG 17-JUL-2002
LOCUS
DEFINITION Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15

```

```

UNORDERED PIECES.
AC127470.1 GI:21886866
HTG: HTGS_PHASE1, HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes
Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Eutheria: Primates: Catarrhini: Homnidae: Pan.
1 (bases 1 to 215795)
REFERENCE
AUTHORS
Antler, N., Antonellis, A., Ayala, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P.,
Lee-Jin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Mastaglio, C., Maskeri, B., Masrlian, S.D.,
McCluskey, J.H., McDowell, J., Paquinigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripp, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Welshby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative.
Unpublished
2 (bases 1 to 215795)
Green, E.D.
Direct Submission
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehgri.nih.gov
----- Project Information
Center project name: cms
Center clone name: 145D13
----- Summary Statistics
Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203629 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-contigs
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
3593: contig of 3593 bp in length
3594 3693: gap of unknown length
3694 8001: contig of 4308 bp in length
8002 8101: gap of unknown length
8102 15639: contig of 7538 bp in length
15640 15739: gap of unknown length
15740 24775: contig of 9036 bp in length
24776 24875: gap of unknown length
24876 30150: gap of unknown length
30151 30250: gap of 5275 bp in length
30251 39212: contig of unknown length
39213 39312: contig of 8962 bp in length
39313 49342: gap of unknown length
49343 49442: contig of 10030 bp in length
49443 60956: gap of unknown length
60956 61056: contig of 11514 bp in length
61057 74520: gap of unknown length
74521 89522: contig of 14902 bp in length

```







Accession	Protein	Length
Db 42506	GGGGCTCTCGCGGGCCCCAGGTCACCTTATGAGGGAGGCGTGTACCTGAAGCT	42447
Oy 155	uasptcLeuValaAngLyAlaLeuAlaLeuArgGysLeuGluGluIupheserAlaThrAl	175
Db 42446	GGACTTGTGTGGATGACGCCCTGGCTTGGCTGGCTGGAGAGTTCTCCGCCAACAC	42387
Oy 175	aAlaserSerProGlyProGlnLeuArgLeuGysGlnValSerGlyLeuLeuProLeuAr	195
Db 42386	GGCGAGCGCCGTGGGCCCCCAGCTCGGCTGCCAACAGTGTGGCGTTCGCCCTCGG	42337
Oy 195	gProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuValaAlaProPh	215
Db 42326	GCCGGGGrCmCCCTGGGGATCCGACCCCTCCCTGGGGCCCATTCACAGCGGCCCTT	42267
Oy 215	eLeuThrTyrTrpHeGlyLeuPheGlnValaHis 225	
Db 42266	CCTTACCTACTTCGGAGCTTCTCCAGGTTAAC 42236	

RESULT	14
LOCUS	AC119115
DEFINITION	AC119115 138792 bp DNA linear HTG 18-JUL-2002 Rattus norvegicus clone CH230-32ON23, *** SEQUENCING IN PROGRESS
ACCESSION	AC119115
VERSION	***, 32 unordered pieces.
KEYWORDS	AC119115 AC119115.2 GI:21746718
SOURCE	HTG; HTGS_PHASE1.
ORGANISM	Norway rat. Rattus norvegicus

## REFERENCE AUTHORS

1. (bases to 138792)

Mizuy, D.M., Adams, C., Adlo-Ochola, B., Al-ossman, F.R., Allen, C., Alsbrooks, S.L., Amerlunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbarella, J., Benton, J., Blmage, K., Blankenburg, K., Bonni, D., Bouch, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N. P., Bukey, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carion, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dattorne, S.R., David, R., Davila, M.L., Davis, C., Day-Carroll, L., Deckerich, D.A., Delaney, K.R., Delgado, O., Degen, A.L., Ding, Y., Dim, H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferriguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gottell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, R., Hart, M., Havlik, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, L.E., Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.B., Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, T., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, Y., Li, Z., Licharge, O., Lieu, C., Liu, T., Liu, W., Louisgeed, H., Lozado, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Matindale, A., Martinez, E., Massey, E., Maxwell, E., Mcloed, M.P., Meadow, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabdat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S., Ogulu, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojupokar, I., Rolfe, M., Ruiz, S., Savery, G., Scherger, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonakke, T., Sparks, A., Stanley, H., Stone, H., Sultun, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vazquez, L., Vera, Y., Villalón, D., Vinson, R., Wang, S., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 138792)
AUTHORS	Worley, R.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 138792)
AUTHORS	Worley, R.C.
TITLE	Direct Submission
JOURNAL	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jul 14, 2002 this sequence version replaced gi:20303440.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GUXG  
Center clone name: CH230-320N23  
Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 112720 bases at least Q40  
Consensus quality: 116666 bases at least Q30  
Consensus quality: 119165 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.bsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.bsc.bcm.tmc.edu/docs/genbank_drift_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 32 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1	1022:	contig of 1022 bp in length
1023	1122:	gap of unknown length
1123	2394:	contig of 1272 bp in length
2395	2494:	gap of unknown length
2495	3663:	contig of 1169 bp in length
3664	3763:	gap of unknown length
3764	4879:	contig of 1116 bp in length
4880	4979:	gap of unknown length
4980	6787:	contig of 1808 bp in length
6788	6882:	gap of unknown length
6888	8192:	contig of 1305 bp in length
8193	8292:	gap of unknown length
8293	9461:	contig of 1169 bp in length
9462	9561:	gap of unknown length
9562	11435:	contig of 1874 bp in length
11436	11535:	gap of unknown length
11536	13535:	contig of 2000 bp in length
13536	13635:	gap of unknown length
13636	17323:	contig of 3688 bp in length
17324	17423:	gap of unknown length
17424	20406:	contig of 2983 bp in length
20407	20506:	gap of unknown length
20507	23445:	contig of 2939 bp in length
23446	23545:	gap of unknown length
23546	25600:	contig of 2055 bp in length
25601	25700:	gap of unknown length
25701	28049:	contig of 2349 bp in length
28050	28149:	gap of unknown length
28150	30474:	contig of 2325 bp in length
30475	30574:	gap of unknown length
30575	35498:	contig of 4924 bp in length
35499	35598:	gap of unknown length



```

* 2413 5841: contig of 3429 bp in length
* 5842 5941: gap of unknown length
* 5942 8435: contig of 2494 bp in length
* 8436 8535: gap of unknown length
* 8536 15799: contig of 7264 bp in length
* 15800 15899: gap of unknown length
* 15900 25224: contig of 9325 bp in length
* 25225 25324: gap of unknown length
* 25325 32504: contig of 7180 bp in length
* 32505 32604: gap of unknown length
* 32605 40970: contig of 8366 bp in length
* 40971 41071: gap of unknown length
* 41072 56590: contig of 15520 bp in length
* 56591 56690: gap of unknown length
* 56691 73770: contig of 17079 bp in length
* 73771 73869: gap of unknown length
* 73870 90859: contig of 16990 bp in length
* 90860 90959: gap of unknown length
* 90960 111428: contig of 20469 bp in length
* 111429 111528: gap of unknown length
* 111529 153553: contig of 42025 bp in length.

```

FEATURES  
Source

```

1. 153553
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_11b="RP42"
1. 2312
/note="assembly_fragment"
2413. 5841
/note="assembly_fragment"
clone_end:17
vector_side:left"
5942. 8435
/note="assembly_fragment"
8536. 15799
/note="assembly_fragment"
15900. 25224
/note="assembly_fragment"
25325. 32504
/note="assembly_fragment"
32605. 40970
/note="assembly_fragment"
41071. 56590
/note="assembly_fragment"
56691. 73769
/note="assembly_fragment"
73870. 90859
/note="assembly_fragment"
90960. 111428
/note="assembly_fragment"
111529. 153553
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

```

BASE COUNT 39178 a 37393 c 36616 g 39259 t 1107 others  
ORIGIN

## Alignment Scores:

```

Pred. No.: 2.81e-27 Length: 153553
Score: 503.00 Matches: 111
Percent Similarity: 58.59% Conservative: 5
Best Local Similarity: 56.06% Mismatches: 17
Query Match: 43.29% Indels: 66
DB: 2 Gaps: 1

```

US-09-245-198A-2 (1-225) x AC126921 (1-153553)

```

OY 93 ProgluGlnaspGlyAlaGlnAlaGlyValaspGlyThrValasergGlyTrpGlnGlnThr 112
DB 26848 CCTGCTCTGTGAACCCACATTCAGGTGTGACGGGACGCTAGTGTGGAGGAGGCC 26907
OY 113 LysIleasnSerserProleuArgTyrAspArgInIleGlyGlnPheThrValIle 132

```

```

DB 26908 AAAATCAACAGCTCCACACCCCTGGCTATGACTCCACAGCCGGGCAATTAGCTCACC 26967
OY 133 ArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 143
DB 26968 CGGGCTGGGCTGTACTACTGCTAGTGTACAGT-AGCCCCACCTGGCTCCACGGGTAAAG 27026
OY 143 143
DB 27027 CCGGAACGTGAAGAGAAAGCTGGCTGGGCTTGGGGGCAAGTTAAAGTGGGAGGG 27086
OY 143 143
DB 27087 GAGCGTGGGCTTGGGCTGAGAGAGAGCCTTGGGCTTAAGACACCTGAGATGAAGCC 27146
OY 144 144
DB 27147 AGGGCCACAGAGGCTGAGCTCCGCCCTCCCTCCCTCCGCCCCACAGTGACATTGATGAG 27206
OY 148 GlyLysAlaValTyrLeuLysIleuAspIleuValAsnGlyValIleuAlaLeuArgCys 167
DB 27207 GGGAAAGCTGTCTACTTAAAGCTGGACTGTGTGATGACACCGCTGGCCCTGGCTGC 27266
OY 168 LeuGlnGluPheSerserAlaThrAlaAlaIleSerserProGlyProGlnLeuArgLeuCysGln 187
DB 27267 CTGGAGGAATTCCTCGCCACCTCGCCGACGCTTCCCTGGGGCCCACTCCGCTCTGCCAA 27326
OY 188 ValSergIleuLeuProleuArgProGlySerserIleuArgIleArgThrLeuProTrp 207
DB 27327 GTGTCAGGCTGTCTTCCGCTCGCGGAGGAGGCTTCCCTCGCGGATCCGACCCCTCCAG 27386
OY 208 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225
DB 27387 ACCCAACTGAAGGCTGCCCTCCCTCTCACTTGTGACTCTTCCAAAGTTTAC 27440

```

Search completed: May 8, 2003, 03:40:19  
Job time : 2022.08 secs

**THIS PAGE BLANK (USPTO)**







PF 04-SEP-1998; 98WO-US18393.  
 XX  
 XX 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 XX Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 XX N-PSDB; AAX23425.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfectected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX Sequence 211 AA;  
 SQ  
 Query Match 93.7%; Score 1089; DB 20; Length 211;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-105;  
 Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 15 LVVSLGSMATLSAQEPSEELTAEDRREPPELNPTQESQDVPELQVPRRSAPKG 74  
 DB 1 LVVSLGSMATLSAQEPSEELTAEDRREPPELNPTQESQDVPELQVPRRSAPKG 60  
 OY 75 RKARPRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSEPLRDROIGFEFTVIRA 134  
 DB 61 RKARPRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSEPLRDROIGFEFTVIRA 120  
 OY 135 GLYLYICQVHDEGKAVYIKLIDLNVGYALRCLDEFSATTAASGPGQIRLCQVSGLLPL 194  
 DB 121 GLYLYICQVHDEGKAVYIKLIDLNVGYALRCLDEFSATTAASGPGQIRLCQVSGLLPL 180  
 OY 195 RFGSSLRIRITLPMAHKAPFLTYGFLQVH 225  
 DB 181 RFGSSLRIRITLPMAHKAPFLTYGFLQVH 211  
 RESULT 4  
 AAM29745  
 ID AAM29745 standard; Protein: 249 AA.  
 XX  
 AC AAM29745;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE TNF related endothelium proliferative agent protein.  
 XX  
 KW TNF; endothelium proliferative agent; TREPA; wound healing; cancer;

KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PE 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 1998-447255/38.  
 DR N-PSDB; AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX Sequence 249 AA;  
 SQ  
 Query Match 87.8%; Score 1020; DB 19; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 OY 2 ISLGLALACGLILLVVSIGSWATLSAQEPSEELTAEDRREPPELNPTQESQDVPEL 61  
 DB 26 LGLGLALACGLILLVVSIGSWATLSAQEPSEELTAEDRREPPELNPTQESQDVPEL 85  
 OY 62 EQLVPRRSAPGRARRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSEPLR 121  
 DB 86 NQLVPRRSAPGRARRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSEPLR 145  
 OY 122 YDRQIGFEFTVIRAGLYLYICQVHDEGKAVYIKLIDLNVGYALRCLDEFSATTAASGPG 181  
 DB 146 YKROIGFEFTVIRAGLYLYICQVHDEGKAVYIKLIDLNVGYALRCLDEFSATTAASGPG 205  
 OY 182 QLRLCQVSGLLPLRPGSSLRIRITLPMAHKAPFLTYGFLQVH 225  
 DB 206 QLRLCQVSGLLPLRPGSSLRIRITLPMAHKAPFLTYGFLQVH 249  
 RESULT 5  
 AAY09369  
 ID AAY09369 standard; Protein: 249 AA.  
 XX  
 AC AAY09369;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Human tumour necrosis factor Apo-3 ligand protein sequence.

XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
 KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
 KW cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09919490-A1.  
 XX  
 PD 22-APR-1999.  
 XX  
 PF 09-OCT-1998; 98WO-US21407.  
 XX  
 PR 17-DEC-1997; 97US-0069862.  
 PR 10-OCT-1997; 97US-0062037.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Masters SA, Pitti R;  
 DR WPI: 1999-287982/24.  
 DR N-PSDB; AAX56000.  
 XX  
 PT New human Apo-3 ligand (a tumour necrosis factor) homologue  
 PS  
 PS Claim 1; Fig 1; 74pp; English.  
 XX  
 CC The present sequence represents a human tumour necrosis factor (TNF)  
 CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
 CC to induce JNK/SAPK-dependent responses in mammalian cells.  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 20; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSEBELTAEDRREPELNPQTEESODVYVFL 61  
 Db 26 LGGLALACGLLLAVVSLGSRASLSAQEPQELVAEEDDPSELNPQTEESODPAPFL 85  
 QY 62 EQLYRPRRSAPKGRKAPRRARIAAHYEVHPRPGDGAQAGVDGVSQWETKINSSPLR 121  
 Db 86 NRLVRPRRSAPKGRKTRARRIAAHYEVHPRPGDGAQAGVDGVSQWETKINSSPLR 145  
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEFSATASSPGP 181  
 Db 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEFSATASSLGP 205  
 QY 182 QLRICQVSGLLPLRPGSSLRIRTLPMANLKAAPLTYFGLEQVH 225  
 Db 206 QLRICQVSGLLALRPGSSLRIRTLPMANLKAAPLTYFGLEQVH 249  
 RESULT 6  
 AAB07526  
 ID AAB07526 standard; protein: 249 AA.  
 XX  
 AC AAB07526;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant human TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200042073-A1.

XX 20-JUL-2000.  
 PD  
 XX 14-JAN-2000; 2000WO-US01044.  
 PF  
 XX 15-JAN-1999; 99US-0116168.  
 PR  
 XX (BIOJ ) BIOGEN INC.  
 XX  
 PI Renner P;  
 DR WPI: 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.  
 XX  
 CC The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune  
 CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 21; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSEBELTAEDRREPELNPQTEESODVYVFL 61  
 Db 26 LGGLALACGLLLAVVSLGSRASLSAQEPQELVAEEDDPSELNPQTEESODPAPFL 85  
 QY 62 EQLYRPRRSAPKGRKAPRRARIAAHYEVHPRPGDGAQAGVDGVSQWETKINSSPLR 121  
 Db 86 NRLVRPRRSAPKGRKTRARRIAAHYEVHPRPGDGAQAGVDGVSQWETKINSSPLR 145  
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEFSATASSPGP 181  
 Db 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEFSATASSLGP 205  
 QY 182 QLRICQVSGLLPLRPGSSLRIRTLPMANLKAAPLTYFGLEQVH 225  
 Db 206 QLRICQVSGLLALRPGSSLRIRTLPMANLKAAPLTYFGLEQVH 249  
 RESULT 7  
 AAY95338  
 ID AAY95338 standard; protein: 249 AA.  
 XX  
 AC AAY95338;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 antitumour protein.  
 XX  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..40

FT Protein /label= signal\_peptide  
 FT 41..249 /label= PRO207  
 FT Modified-site 27..33 /note= "N-myristoylation"  
 FT Modified-site 29..35 /note= "N-myristoylation"  
 FT Modified-site 36..42 /note= "N-myristoylation"  
 FT Modified-site 45..51 /note= "N-myristoylation"  
 FT Modified-site 118..124 /note= "N-myristoylation"  
 FT Modified-site 121..127 /note= "N-myristoylation"  
 FT Modified-site 125..131 /note= "N-myristoylation"  
 FT Modified-site 128..134 /note= "N-myristoylation"  
 FT Modified-site 139..143 /note= "N-myristoylation"  
 FT Modified-site 10..14 /note= "Asn is N-glycosylated"  
 FT Modified-site 97..101 /note= "amidation"  
 FT Modified-site 24..35 /note= "amidation"  
 FT Peptide /note= "prokaryotic membrane lipoprotein lipid"  
 FT WO200037638-A2.  
 XX 29-JUN-2000.  
 PD 02-DEC-1999; 99WO-US28565.  
 XX 22-DEC-1998; 98US-013296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX (GENH ) GENENTECH INC.  
 PA Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pictl RM, Wood WI;  
 XX N-PSDB; AAA49717.  
 DR WPI: 2000-442668/38.  
 XX Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or  
 PT PRO866  
 XX Claim 19; Fig 4; 172pp; English.  
 PS  
 XX The present sequence is that of human antitumour protein PRO207,  
 CC as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207  
 CC shows amino acid sequence identity to tumour necrosis factor family  
 CC members, especially human lymphotoxin-beta (23.4%) and human CD40  
 CC ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting  
 CC the growth of a tumour cell comprises exposing the tumour cell  
 CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO328, PRO301,  
 CC PRO526, PRO362, PRO509 or PRO866 (see AAY95337-49), their  
 CC agonists or chimeric polypeptides incorporating them. The tumour  
 CC is especially a cancer selected from breast, ovarian, renal,  
 CC colorectal, uterine, prostate, lung, bladder and central nervous  
 CC system cancer, melanoma and leukaemia. Methods for the recombinant  
 CC expression of the antitumour proteins are also provided.

SQL Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 21; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLVVSLGSMATLSAOPSEETADREPPPLNPTQTESQDPVFL 61  
 DB 26 LGLGLALACGLLVVSLGSMATLSAOPSEETADREPPPLNPTQTESQDPVFL 85  
 QY 62 EQLVPRRSAPKGRKARRRRAIAHYEVHPRGQGAQAGVDGTVSGMEETKINSSEPLR 121  
 DB 86 NKLVRPRRSAPKGRKARRRRAIAHYEVHPRGQGAQAGVDGTVSGMEETKINSSEPLR 145  
 QY 122 YDRQIGFEFTVIRAGLYLYICQVHFDEGKAVYIKLDDLNGVIALCLDEFSSTAASSRCP 181  
 DB 146 YNRQIGFEFTVIRAGLYLYICQVHFDEGKAVYIKLDDLNGVIALCLDEFSSTAASSRCP 205  
 QY 182 QLRLQVSGILPLRPGSSLRIRTLPMALUKAPFLTYFGLPQVH 225  
 DB 206 QLRLQVSGILALRRGSSLRIRTLPMALUKAPFLTYFGLPQVH 249  
 RESULT 8  
 ID AAE00891 standard; Protein: 249 AA.  
 AC AAE00891;  
 DE 04-JUL-2001 (first entry)  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vuneratery.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 98..249 /label= Extracellular\_domain  
 FT US6207642-B1.  
 PN 27-MAR-2001.  
 PD 26-JUN-1998; 98US-0105343.  
 XX 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX (ABBO ) ABBOTT LAB.  
 PA Wiley SR;  
 PI WPI: 2001-280760/29.  
 DR N-PSDB; AAD04350.  
 XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT -  
 XX Claim 1; Column 75-76; 53pp; English.  
 PS  
 XX The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.

CC The present amino acid sequence is clone ID #690050 human TREPA.

50 Sequence 249 AA;

Query Match 87.88; Score 1020; DB 22; Length 249;

Best Local Similarity 88.88; Pred. No. 9.5e-98;

Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY	2	LSLGLATACGILLVVYVSLGSMATLSAOEPOEELTVEDDRREPENLAPOTEEESODVYFL	61
Db	26	LGIGLATACGILLVVYVSLGSRASLSAOEPOEELVAEEDDSELPOTEEESODPAPFL	85
QY	62	EOLVPRPRSAFGKARKAPRRRAIAAHVEVHPRAPGOGAAGVDGVSQMEETKINSPPLR	121
Db	86	NRLVPRPRSAFGKARKTARRRIAIAHVEVHPRAPGOGAAGVDGVSQMEETKINSPPLR	145
QY	122	YDRQIGFEFTVIRAGLLTYLYLCOVHDEGRKAYLIKIDLLVNGVLALRCLFEESATAPSSGP	181
Db	146	YNROIGFEFTVIRAGLLTYLYLCOVHDEGRKAYLIKIDLLVDGVLALRCLFEESATAPSSGP	205
QY	182	QLRLCOVSGLLPRLPGSSSLRRTLPMAHLKAAAPLLTYFGFLQVH	225
Db	206	QLRLCOVSGLLALRPGSSSLRRTLPMAMHLKAAAPLLTYFGFLQVH	249

RESULT 9  
AAU86129  
ID AAU86129 standard; Protein; 249 AA

AC AAU86129;

DT 15-JUL-2002 (first entry)

DE Human PRO207 polypeptide.

KM Human, PPO; benign tumour; malignant tumour; lymphoid malignancy;  
 KW leukaemia; neuronal disorder; stromal disorder; blastocoeal disorder  
 KM Inflammatory disorder; immune disorder; angiogenic disorder;  
 KM cytosstatic; neuroprotective.  
 XS Homo sapiens.  
 XS

OS Homo sapiens.

PN WO200153486-A1

PD 26-JUL-2001

11-FEB-2000; 2000WO-US03565.

PR	08-MAR-1999;	99WO-US05028.
PR	11-MAR-1999;	99US-123972P.
PR	11-MAY-1999;	99US-133459P.
PR	02-JUN-1999;	99WO-US12252.
PR	22-JUN-1999;	99US-140650P.
PR	22-JUN-1999;	99US-140653P.
PR	20-JUL-1999;	99US-144758P.
PR	26-JUL-1999;	99US-145698P.
PR	28-JUL-1999;	99US-146322P.
PR	17-AUG-1999;	99US-149395P.
PR	31-AUG-1999;	99US-151689P.
PR	01-SEP-1999;	99WO-US20111.
PR	15-SEP-1999;	99WO-US21090.
PR	30-NOV-1999;	99WO-US28313.
PR	01-DEC-1999;	99WO-US28301.
PR	01-DEC-1999;	99WO-US28634.
PR	05-JAN-2000;	2000WO-US00219.

PA (GETH ) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

PI Watanabe CK, Wood WI;

DR WPI: 2002-205567/26.

DR WPI: 2002-205567/26.

DR N-PSDB; ABK40255

PT Thirty five nucleic acids encoding PRO polypeptides, useful for  
PT treating benign or malignant tumours, leukaemias and lymphoid  
PT malignancies, inflammatory, angiogenic and immunologic disorders  
XX  
PS Claim 61; F1g 4; 302pp; English.

CC The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The  
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular  
CC macrophasal, stromal and blastocoellic disorders, inflammatory, immune  
CC and angiogenic disorders. The polynucleotide sequences are also  
CC useful in gene therapy. AA086128-AA086152 represent the human PRO  
CC polypeptides of the invention.

**SQ**      **Sequence**      **249**      **AA;**

Query Match	87.88;	Score 1020;	DB 23;	Length 249;
-------------	--------	-------------	--------	-------------

Best Local Similarity 88.8%; Pred. No. 9.5e-98;

Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY	2	LSGIALACIGLLLVYSLGSMALTSQEPSEOEELTEDREPELNPQSESDVYPL	61
Db	26	LGLGIALACIGLLLVYSLGSRASLSAQEPSEOEELVVEEDDPSELNPQSESDPAPPL	85
QY	62	EOLVPRPSAPKGRKARPRRAIAAHYEVPHPGDDGAQVGDGVSGWEETKINSSPPLR	121
Db	86	NRLVPRPSAPKGRKTRARRRAIAAHYEVHPHPGDDGAQVGDGVSGWEKARINSSPPLR	145
QY	122	YDRQIGETVTRAGIYYLYYCQVHDEGKAAVYLKLDLLVNGVTLARCLSEFSATAASSRP	181
Db	146	YNRQIGETVTRAGIYYLYYCQVHDEGKAAVYLKLDLLVDVTLARCLSEFSATAASSLGP	205
QY	182	QLRLQVSGLLPLRPGSSLRTPLMAHLKAAPLYTFGLFQVH	225
Db	206	QLRLQVSGLLPLRPGSSLRTPLMAHLKAAPLYTFGLFQVH	249

RESULT 10	
AAW47525	
ID	AAW47525 standard; Protein; 284 AA

AC AAW47525;

DT 21-JUL-1998 (first entry)

DE Homo sapiens tumour necrosis factor related ligand (TNFRL)

KW TRELL; tumour necrosis factor related ligand; tnf; treatment;  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection.

OS Homo sapiens.

PN W09805783-A1.

PD 12-FEB-1998.

PF 07-AUG-1997; 97WO-US13945.

PR 18-MAR-1997; 97US-0040820

PR 18-OCT-1996; 96US-0028515.

PA (BIOJ ) BIOGEN INC.

PA (UYGE-) UNIV GENEVA FACULTY MEDICINE

PI Browning JL, Chicheportliche Y;

[illegible]



ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 rubrosis; uveitis; macular degeneration; arthritis; rheumatism;  
 corneal graft neovascularisation; psoriasis; metastatic condition;  
 malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 preneoplastic condition; myocardial angiogenesis; wound granulation;  
 scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 fusion protein.

Homo sapiens.  
 Synthetic.

MO200145730-A2.

28-JUN-2001.

19-DEC-2000; 2000MO-US34755.

20-DEC-1999; 98US-0172878.

10-MAY-2000; 2000US-0203347.

(IMV ) IMMUNEX CORP.

Wiley SR;

WPI: 2001-417975/44.

N-PSDB: AAS03964.

Modulating angiogenesis in a mammal for treating diseases mediated by  
 angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 peripheral tissue, by administering antagonist or agonist of TWEAK  
 receptor

Example 1; Page 41; 46pp; English.

The sequence represents a fusion protein encoded by the  
 expression vector pDC409-L2-TWEAK. The fusion protein comprises a  
 growth hormone leader, a leucine zipper multimerisation domain, and  
 the extracellular domain of human TWEAK. The fusion protein was  
 used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 clones from a COS cell human cDNA library. The TWEAK protein is  
 a member of the tumour necrosis factor (TNF) family and induces  
 angiogenesis. TWEAKR may therefore be used to screen for and  
 develop TWEAKR agonists and antagonists for the modulation of  
 angiogenesis, to be used in the treatment and diagnosis of human disease.  
 The disorders mediated by angiogenesis include ocular disorders  
 characterised by ocular neovascularisation such as diabetic retinopathy,  
 neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 retrolental fibroplasia, rubrosis, uveitis, macular degeneration and  
 arthritis, rheumatism and psoriasis. Other treatable diseases such as  
 malignant and metastatic conditions such as sarcomas and carcinomas,  
 benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 plaque neovascularisation, telangiectasia, wound granulation, coronary  
 atherosclerosis, peripheral atherosclerosis and ischaemia.

Sequence 273 AA;

Query Match 81.88; Score 951; DB 22; Length 273;

Best Local Similarity 88.98; Pred. NO. 1.6e-90;

Matches 184; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 19 SLGSMATLSAQPSEELTAEDREPPPELNPQTESQDVVFLQOLVPRRSAPKGRAR 78  
 DB 67 SLGSRASLSAQEPADQELVAEDDPSPELNQTESQDPAPFLNRLVPRRSAPKGRKTR 126  
 QY 79 PRRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGETVIRAGLY 138  
 DB 127 ARRAlAAHYEHVPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGETVIRAGLY 186

QY 139 LYCQVHFDEGKAVYIKLIDLNVGVALRCLREFSATMASSPGPOLRLCOVSGILPLRPS 198  
 DB 187 LYCQVHFDEGKAVYIKLIDLNVGVALRCLREFSATMASSPGPOLRLCOVSGILPLRPS 246  
 QY 199 SIRIRLPAHAKAPFLYFGLFOVH 225  
 DB 247 SIRIRLPAHAKAPFLYFGLFOVH 273

RESULT 13

AAW29746

ID AAW29746 standard; Protein; 189 AA.

XX AAW29746;

AC 27-OCT-1998 (first entry)

DE TNF related endothelium proliferative agent protein 2.

XX TNF; endothelium proliferative agent; TREPA; wound healing; cancer;

KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

XX Homo sapiens.

OS W09835061-A2.

PN 13-AUG-1998.

PD 12-FEB-1998; 98WO-US02859.

PF 10-FEB-1998; 98US-0021706.

XX 12-FEB-1997; 97US-0798692.

PR (ABBO ) ABBOTT LAB.

PA WILEY SR;

PI WPI: 1998-447255/38.

DR Detecting nucleic acid encoding TREPA - useful for diagnosis and

XX treatment of autoimmune disease, tumours and inflammation

PT Claim 16; Page 125-6; 142pp; English.

The TNF-related endothelium proliferative agent (TREPA), or its  
 activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 promote wound healing or tissue grafting, by promoting vascularisation,  
 also to induce apoptosis for treating cancer and eliminating autoreactive  
 T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 TREPA peptides can also be used to target cytotoxic agents or for  
 affinity isolation of the corresponding receptor, the nucleic acid for  
 which can be used to transform tumour cells to render them more  
 responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).

Sequence 189 AA;

Query Match 63.98; Score 742; DB 19; Length 189;

Best Local Similarity 92.88; Pred. No. 5.7e-69;

Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 73 KGRARPRRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGETVY 132  
 DB 37 KGRTRARRRGIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGETVY 96  
 QY 133 RAGLYLYLCQVHFDEGKAVYIKLIDLNVGVALRCLREFSATMASSPGPOLRLCOVSGIL 192  
 DB 97 RAGLYLYLCQVHFDEGKAVYIKLIDLNVGVALRCLREFSATMASSPGPOLRLCOVSGIL 156



```
OY 193 PLRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||
DB 157 ALRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 189

RESULT 14
AAE00892
ID AAE00892 standard; Protein; 189 AA.
XX
AC AAE00892;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human ULAflag TREPA soluble construct.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;
KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;
KW vulnery; HUVEC; human umbilical vein endothelial cell; ULAflag.
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-0105343.
XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
CC -
CC Example 2; Column 75-78; 53pp; English.
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts. The present
CC amino acid sequence is human ULAflag TREPA soluble construct. This
CC sequence which is a biologically active molecule is capable of inducing
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.
XX
SQ Sequence 189 AA;
XX
Query Match 63.9%; Score 742; DB 22; Length 189;
Best Local Similarity 92.8%; Pred. No. 5.7e-69;
Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 73 KGRKARPRRAIAHYEVHPRPGODGAQAGVDGTVSGMEETKINSSPLRYDRQIGFTYI 132
      |||||
DB 37 KGRKTRARGIAHYEVHPRPGODGAQAGVDGTVSGMEARINSSPLRYNQGIFTYIT 96

OY 133 RAGLYLYLCQVHFDEGKAVYLLKLDLVNGVIALRCLEEFSSATPAASPQQLRCQVSGLL 192
      |||||
DB 97 RAGLYLYLCQVHFDEGKAVYLLKLDLVNGVIALRCLEEFSSATPAASPQQLRCQVSGLL 156

OY 193 PLRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||
DB 157 ALRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 189
```

```
RESULT 15
AAE00895
ID AAE00895 standard; Protein; 146 AA.
XX
AC AAE00895;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) fragment.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnery.
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-0105343.
XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
CC -
CC Example 14; Fig.1; 53pp; English.
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts.
CC The present amino acid sequence is human TREPA fragment.
XX
SQ Sequence 146 AA;
XX
Query Match 62.0%; Score 721; DB 22; Length 146;
Best Local Similarity 94.5%; Pred. No. 6.1e-67;
Matches 138; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 80 RRAIAAHYEVHPRPGODGAQAGVDGTVSGMEETKINSSPLRYDRQIGFTYIRAGLYYL 139
      |||||
DB 1 RRAIAAHYEVHPRPGODGAQAGVDGTVSGMEARINSSPLRYNQGIFTYIRAGLYYL 60

OY 140 YQVHFDEGKAVYLLKLDLVNGVIALRCLEEFSSATPAASPQQLRCQVSGLLPLRPGSS 199
      |||||
DB 61 YQVHFDEGKAVYLLKLDLVNGVIALRCLEEFSSATPAASPQQLRCQVSGLLALRPGSS 120

OY 200 LRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||
DB 121 LRIRLTPMAHLKAAPFLTYFGLFOVH 146
```

Search completed: May 8, 2003, 02:19:41  
Job time : 29.7328 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:39 ; Search time 1561.93 Seconds  
(without alignments)  
12110.869 Million cell updates/sec

Title: US-09-245-198a-1  
Perfect score: 1168  
Sequence: 1 ggtgtgcgacctggcctg.....ataatcatgtctcttc 1168

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapept 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883.8	75.7	1033	11	AK020909 Mus musculus
2	625.8	53.6	918	12	BF577781 BF577781
3	519.4	44.5	731	13	BI871711 BI871711
4	507.4	43.4	728	13	BI870393 BI870393
5	504	43.2	561	10	AM763237 ur70d09.y
6	497	42.6	533	10	BE628951 BE628951

7	489.6	41.9	650	12	BG404836	BG404836 602240160
8	488.8	41.8	687	14	BO208433	BO208433 UT-R-EP0
9	481.2	41.2	584	10	AM917574	AM917574 EST348878
10	480.4	41.1	777	13	BI819200	BI819200 603034614
11	473	40.5	834	13	BI766766	BI766766 603056866
12	467.8	40.1	948	14	BO707185	BO707185 AGENCOURT
13	456.4	39.1	828	13	BI596681	BI596681 603243254
14	437.8	37.5	471	9	AA221610	AA221610 my18d09.y
15	421.6	36.1	963	14	BO671259	BO671259 AGENCOURT
16	419.2	35.9	940	14	BO884231	BO884231 AGENCOURT
17	406.4	34.8	1071	14	BM921213	BM921213 AGENCOURT
18	390.6	33.4	445	9	AA870722	AA870722 vq25q07.y
19	366.6	31.4	951	14	BO674188	BO674188 AGENCOURT
20	360	30.8	360	10	BE654876	BE654876 UT-M-BHO-
21	353	30.2	405	9	AT854476	AT854476 UT-M-BHO-
22	331	28.3	367	13	BI111534	BI111534 602895883
23	325.4	27.9	376	9	AA792068	AA792068 vns6d09.y
24	309	26.5	474	13	BI965174	BI965174 1d34e07.y
25	306	26.2	315	12	BE466521	BE466521 UT-M-CGP
26	299.4	25.6	538	12	BF821434	BF821434 MR1-RT003
27	297.2	25.4	400	9	AT152313	AT152313 ud87h02.y
28	292.8	25.1	493	10	BE307031	BE307031 601087888
29	289.6	24.8	433	12	BF283688	BF283688 EST448279
30	289.4	24.8	542	12	BF041509	BF041509 BP250025B
31	285.2	24.4	443	12	BG378802	BG378802 UT-R-CV1-
32	284.4	24.3	785	13	BI762908	BI762908 603047966
33	283.8	24.3	404	12	BF044430	BF044430 BP250013A
34	278.4	23.8	897	13	BI730298	BI730298 603350276
35	274.4	23.5	412	9	AT101416	AT101416 EST204867
36	271	23.2	581	13	BI738634	BI738634 603358846
37	270.4	23.2	402	12	BF410871	BF410871 UT-R-CNO-
38	270.2	23.1	292	10	BB268794	BB268794 BB268794
39	258.2	22.1	558	13	BM484863	BM484863 538865 MA
40	255.2	21.8	394	12	BG376757	BG376757 UT-R-CU0-
41	246.4	21.1	298	14	BM688946	BM688946 UT-E-CQ1-
42	243.6	20.9	371	13	BI300370	BI300370 UT-R-CV2-
43	242.4	20.8	441	13	BI967060	BI967060 1d27h07.y
44	242.2	20.7	374	9	AA800970	AA800970 EST190467
45	240	20.5	240	10	AM764050	AM764050 ur70d09.x

## ALIGNMENTS

RESULT 1	AK020909	1033 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930030D13:tumor necrosis factor (ligand) superfamily, member 12, full insert sequence.				
DEFINITION	AK020909				
ACCESSION	AK020909.1	GI:12861640			
VERSION	AK020909.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult retina cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279283				
PubMed	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20493374				
PubMed	11042159				



D	b		422	CAGATCTACGTAGGGGCGTTTGCTGTCCCAAGATTTCCTTAAGTTCCTGCCGCGCTCCAGAGGC	481
Oy		728	ATCACACACACTCCTCCATCCCACCACCCTCCACTCTCCACCCTCGCTGCTCTTGCTCCAG	787	
D	b	482	ATCACACACACTTNTCATCCCCACCCCCACTCTCCACCCTTAGCTGCTCTTGCTCCAG	541	
Oy		788	TCTCGT-CTGCGTCAAAGGAGCAGAGCTTTGTCATAG-TTTCCAATTCCACAGAGCT	845	
D	b	542	TCTGTCCTCTGCTCAAAGGAGCAGAGCTTTGTCATAGTTTTCCATTCACAGAGCT	601	
Oy		846	ATCTTGCTCTTC-TTAAACATCCCATCCACACACAATACTACCTACCTAGCTCCAAA	904	
D	b	602	ATCCTGTGCTCTCTTTAACAATCCATCCACACAGAATATCAGCTACTAGCTCCAAA	661	
Oy		905	GCCCCCTACTATATCCCTGAAGTCCACCACCTACACCCGACACAGCTGTATTAGACTTTGT	964	
D	b	662	GCCCCCTACTATATCCCTGAAGTCCACCACCTACACCCGACACAGCTGTATTAGACTTTGT	721	
Oy		965	GCACAGGCACTGAGATGGGTGGACCTGGTGGGAGGAGAACAGAACACTGGGAGCTAGG	1024	
D	b	722	GCACAGGCACTGAGATGGGTGGACCTGGTGGGAGGAGAACAGAACACTGGGAGCTAGG	781	
Oy		1025	CCAGAACTTCCCACTGTGAGGGGGGAGAGAGTGGGAGAACTCTCCCTGGATCCCTGT	1084	
D	b	782	CCAGAACTTCCCACTGTGAGGGGGGAGAGAGTGGGAGAACTCTCCCTGGATCCCTGT	841	
Oy		1085	GGATTTTGAAAAAGATCTATTTTTATTTATTTATTTATTTGACAAAATGTTAAATGGATTTAA	1144	
D	b	842	GGATTTTGAAAAAGATCTATTTTTATTTATTTATTTATTTGACAAAATGTTAAATGGATTTAA	901	
Oy		1145	GAGATTAATCATGATTTCTCTC	1168	
D	b	902	GAGATTAATCATGATTTCTCTC	925	
RESULT_2					
LOCUS		BFS77781	918 bp	mRNA	linear EST 12-DEC-2000
DEFINITION		602092080F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206595 5'			
ACCESSION		BFS77781			
VERSION		BFS77781.1	GI:11651493		
KEYWORDS		EST.			
SOURCE		house mouse.			
ORGANISM		Mus musculus			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE		1 (bases 1 to 918)			
JOURNAL		NIH-MDC http://mhc.nci.nih.gov/.			
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)			
		Contact: Robert Strausberg, Ph.D. Email: cgabs@mail.nih.gov Tissue Procurement: Jeffrey F. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLM9767 row: p column: 20 High quality sequence start: 17 High quality sequence stop: 724. Location/Qualifiers 1..918 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4206595" /_lab_host="NCI CGAP Co24" /_lab_host="DHIOB (T1 phage-resistant)" /note="Organ: colon; Vector: pCMV-Sport6; Site:1; NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT."			
FEATURES					
SOURCE					

		Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI/CGAP Library."
BASE COUNT	153 a 292 c 268 g 204 t	1 others
ORIGIN		
Query Match	53.6%; Score 625.8; DB 12;	Length 918;
Best Local Similarity	99.2%; Pred. No. 7.2e-152;	
Matches 660;	Conservative 0; Mismatches 2;	Indels 3; Gaps 3;
OY	22 GCTGGCCTGCTTGGCCCTCTGTGTGTGTGTACAGCTGGGGAGCTGGGCACAGCCTGTC	81
Dd	1 GCTGGCCTGCTTGGCCCTCTGTGTGTGTGTACAGCTGGGGAGCTGGGCACAGCCTGTC	60
OY	82 TGCCAGAGAGCCTTTCAAGAGAGACTCATACAGACAGACGCCCGGAGACCCTTGAACT	141
Dd	61 TGCCAGAGAGCCTTTCAAGAGAGACTCATACAGACAGACGCCCGGAGACCCTTGAACT	120
OY	142 GAATCCCCAGACAGAGAAAGCCAGAGTAGTGTTACTCTTCTTGAACAACATGATGTCGGGC	201
Dd	121 GAATCCCCAGACAGAGAAAGCCAGAGTAGTGTTACTCTTCTTGAACAACATGATGTCGGGC	180
OY	202 TCGAAGAAGTGCTCTTAAGGCGCGAAGGCGCGCCTCGCCAGAGCTATTGACGCCATTYA	261
Dd	181 TCGAAGAAGTGCTCTTAAGGCGCGAAGGCGCGC - CTGGCGAGGTATTGACGCCATTYA	239
OY	262 TGAGTTATCTCTGGCCAGACAGAGATGAGACACACAGCAGSTGTGATGGACAGTGAG	321
Dd	240 TGAGTTATCTCTGGCCAGACAGAGATGAGACACACAGCAGSTGTGATGGAGACAGTGAG	299
OY	322 TGGCTGGGAAGAGACCAAATAACAGCTCCAGCCCTCTGCGCTAGCGACCGCCAGATGAG	381
Dd	300 TGGCTGGGAAGAGACCAAATAACAGCTCCAGCCCTCTGCGCTAGCGACCGCCAGATGAG	359
OY	382 GGAATTACAGTCATCAGAGGCTGGGCTCTACTACCTGTACTGTACAGTGCATTTGATGA	441
Dd	360 GGAATTACAGTCATCAGAGGCTGGGCTCTACTACCTGTACTGTACAGTGCATTTGATGA	419
OY	442 GGGAAAAGCTGTCTACTGAGAGCTGTGACTGTGTGAAAGGTGTGCTGGCCCTGCGCTG	501
Dd	420 GGGAAAAGCTGTCTACTGAGAGCTGTGACTGTGTGAAAGGTGTGCTGGCCCTGCGCTG	479
OY	502 CCTGGAAGAATTTCTAGCGCACAGACAGCAAGCTCTCTGGGCCCCAGCTCCCTTTTGCCA	561
Dd	480 CCTGGAAGAATTTCTAGCGCACAGACAGCAAGCTCTCTGGGCCCCAGCTCCCTTTTGCCA	539
OY	562 GGTCGTGGGCTGTGGCGCTCGGACGAGAGGCTTCCCTTGGATCCGACCCCTCCCTG	621
Dd	540 GGTCGTGGGCTGTGGCGCTCGGACGAGAGGCTTCCCTTGGATCCGACCCCTCCCTG	598
OY	622 GGCATCATTTAAGCTGCCGCCCTTCTTAACCTACTTGTGACACTTTTCAAGTTCACTGAG	681
Dd	599 GGCATCATTTAAGCTG - CCCCCTCTTAACCTTGTGACACTTTTCAAGTTCACTGAG	657
OY	682 GGCTT 686	
Dd	658 GGCTT 662	
RESULT 3		
B1871711	B1871711	731 bp mRNA linear EST 11-OCT-2001
LOCUS	603395825P1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5'	
DEFINITION	mRNA sequence.	
ACCESSION	B1871711	
VERSION	B1871711.1 GI:16045386	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
TAXID	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.	
DATE	1 (bases 1 to 731)	
PROJECT	National Institutes of Health, Mammalian Gene Collection (MGC)	



```

/strain="129.C57BL/6J.FVB/N"
/db_xref="taxon:10090"
/clone_xref="IMAGE:3155693"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"

```

RESULT 6	533 bp	mRNA	linear	EST 25-AUG-2000
LOCUS	BE628951			
DEFINITION	mus30c03.y1 Soares_mammary_gland_NIMMG	Mus musculus	cDNA clone	
	IMAGE:3373444.5	similar to	IR:054907	054907 TNF-RELATED WEAK
	INDUCER OF APOPTOSIS ;			mRNA sequence.
ACCESSION	BE628951			
VERSION	BE628951.1	GI:9911639		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 533)			
AUTHORS	NCI-CCGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
	Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D.			
	Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a>			
	This clone is available royalty-free through LLNL ; contact the			
	IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information.			
	GI:1083048			

Seq primer: -40RP from Gibco  
High quality sequence stop: 437.  
Location/Qualifiers

## FEATURES

source

1..533

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:3373444"

/clone\_lib="Soares\_mammary\_gland\_NKLMG"

/sex="female (lactating)"

/tissue\_type="mammary gland"

/lab\_host="DH10B"

/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker: 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 91 a 206 c 105 g 131 t  
ORIGIN

Query Match 42.6%; Score 497; DB 10; Length 533;  
Best Local Similarity 99.4%; Pred. No. 2e-118;

Matches 530; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

OY 481 CGGTGCGTGGCCGCGCGTGGAGAAATTCACGACAGACAGAGCTGCTGG 540  
DB 1 CGGTGCGTGGCCGCGCGTGGAGAAATTCACGACAGACAGAGCTGCTGG 50  
OY 541 GCCCAGCTCCGTTGTGTCAGAGTGTGCGGCTGTTGCGGCGGCGGCTCCCT 600  
DB 61 GCCCAGCTCCGTTGTGTCAGAGTGTGCGGCTGTTGCGGCGGCGGCTCCCT 120  
OY 601 TCGGATCCGACCCCTCCCTGGGCTCATTAAAGCTGCGCCCTCTTAACCTATTGG 660  
DB 121 TCGGATCCGACCCCTCCCTGGGCTCATTAAAGCTGCGCCCTCTTAACCTATTGG 180  
OY 661 ACTCTTCAAGTTCACAGAGGCGCTGCTCCAGATTCCTTAACTTCCCTGGCTC 720  
DB 181 ACTCTTCAAGTTCACAGAGGCGCTGCTCCAGATTCCTTAACTTCCCTGGCTC 240  
OY 721 CAGGAGCATCAACACACCTCCCTACCCACCCCACTGCTCCGCTGCTGCT 780  
DB 241 CAGGAGCATCAACACACCTCCCTACCCACCCCACTGCTCCGCTGCTGCT 300  
OY 781 GGTCCAGTCTGT -CTTCTCTCAAGAGCAGCAGAGCTTGTACATG -TTCCATTCCA 838  
DB 301 GGTCCAGTCTGTCTCTCTCTCAAGAGCAGCAGAGCTTGTACATGTTTCCATTCCA 360  
OY 839 CAGAGTATCCCTGCTCTC -TTACATCCCATCCACACCAACTATCCTACTAGC 897  
DB 361 CAGAGTATCCCTGCTCTCTCTTAAATCCATCCACACCAACTATCCTACTAGC 420  
OY 898 TCCCAAGCCCTTACTTATCCCTGACTCCCACTCCCACTCCAGCAGCTGTTATATG 957  
DB 421 TCCCAAGCCCTTACTTATCCCTGACTCCCACTCCCACTCCAGCAGCTGTTATATG 480  
OY 958 ACTTGTGACACGAGCAGTGAATGGCTGAGACCTGTGTGAGAGAGGACAG 1010  
DB 481 ACTTGTGACACGAGCAGTGAATGGCTGAGACCTGTGTGAGAGAGGACAG 533

RESULT 7 650 bp mRNA linear EST 12-MAR-2001  
LOCUS BG404836  
DEFINITION 602420160P1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:4527038 5',  
ACCESSION BG404836  
VERSION BG404836.1 GI:13298284  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: L14M10435 row: h column: 15  
High quality sequence stop: 468.  
Location/Qualifiers

## FEATURES

source

1..650

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:4527038"

/clone\_lib="NIH\_MGC\_94"

/tissue\_type="retina"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH-MGC Library."

BASE COUNT 157 a 194 c 149 g 150 t  
ORIGIN

Query Match 41.9%; Score 489.6; DB 12; Length 650;  
Best Local Similarity 92.0%; Pred. No. 1.9e-116;

Matches 550; Conservative 0; Mismatches 44; Indels 4; Gaps 3;

OY 291 GAGCAAGCAGGTGTGATGGAGACAGTGTGCTGGAGAGACCAAAATCAAGCT 350  
DB 25 GAGCAAGCAGGTGTGATGGAGACAGTGTGCTGGAGAGACCAAAATCAAGCT 84  
OY 351 CGAGCCCTGCGCTGACAGCCGCGAGATGGGGAATTTACAGTATAGGCTGGCTCT 410  
DB 85 CGAGCCCTGCGCTGACAGCCGCGAGATGGGGAATTTACAGTATAGGCTGGCTCT 144  
OY 411 ACTACCTGACTGACGTGACATTTGATGAGGAAGAGCTGTACCTGAAGCTGACT 470  
DB 145 ACTACCTGACTGACGTGACATTTGATGAGGAAGAGCTGTACCTGAAGCTGACT 204  
OY 471 TGTGTGTAACGCTGTGCTGCGCTGCGCTGCGCTGCGGAAGAAATTCACAGCAGCAA 530  
DB 205 TGTGTGTAACGCTGTGCTGCGCTGCGCTGCGCTGCGGAAGAAATTCACAGCAGCAA 264  
OY 531 GCTCTCGGGGCGCCAGCTGCTGTGCGCAGGCTGCTGGGCTGCTGCTGCTGCTGCT 590  
DB 265 GCTCTCGGGGCGCCAGCTGCTGTGCGCAGGCTGCTGGGCTGCTGCTGCTGCTGCT 323  
OY 591 GGTCTCCCTTGGATCGGACCGCTCCCTGGGCTGCTTAAAGCTGCGCCCTTCTAA 650  
DB 324 GGTCTCCCTTGGATCGGACCGCTCCCTGGGCTGCTTAAAGCTGCGCCCTTCTAA 383  
OY 651 CTTACTTTGAGCTTTTCAAGTTCAGTGAAGGCGCTGCTGCTGCTGCTTAAACTT 710  
DB 384 CTTACTTTGAGCTTTTCAAGTTCAGTGAAGGCGCTGCTGCTGCTGCTTAAACTT 443  
OY 711 TCCCTGGCTCAGAGAGATACCACTCCCTGACCCCACTCCCACTCCCACTCCCTG 770  
DB 444 TCCCTGGCTCAGAGAGATACCACTCCCTGACCCCACTCCCACTCCCACTCCCTG 503  
OY 771 GCTGCTCTTGTGTCAGTCTGT -CTCTCTCAAGAGCAGCAGAGCTTTCACATGTT 829  
DB 504 GCTGATCTTGTGTCAGTCTGTCTCTCTCAAAAGGACAGACAGCTTGTACATGTA 563



OY 830 TCATTCACAGACGTATCTTGGCTCTTAAATCCATCCACCACTATACCA 887  
 Db 564 ATCAAT--CCAAAGATATCTAGATCTACTTAATAGCCATGCAAAACATATACACA 619  
 RESULT 8  
 LOCUS B0208433/c  
 DEFINITION B0208433 687 bp mRNA linear EST 02-MAY-2002  
 UI-R-EP0-coc-1-22-0-UI s1 UI-R-EP0 Rattus norvegicus cDNA clone  
 UI-R-EP0-coc-1-22-0-UI 3', mRNA sequence.  
 ACCESSION B0208433  
 VERSION B0208433.1 GI:20424898  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 687)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 9/704447  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized duodenum library cDNA library Preparation: M.B. Soares  
 Lab Clone distribution: clones will be available through Research  
 Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..687  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-EP0-coc-1-22-0-UI"  
 /clone\_id="UI-R-EP0"  
 /dev\_stage="ADULT"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pUT3D-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not I; Site 2: Eco RI; UI-R-EP0 is a  
 subtracted cDNA library constructed according to Bonaldo,  
 Lennon and Soares, Genome Research, 6:791-806, 1996. First  
 strand cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pUT3D-Pac vector. The oligonucleotide  
 used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dT)18 tail. The sequence tags for this  
 library are: distal colon, GAAGTGTCC; osteoblast,  
 AAGATATCAA; cell line R3327 5A, GGACTAGATC; cell line  
 R3327 5P, CACGTGAGAT; duodenum, TGCTTCAT; prostate,  
 CCAGG.  
 TAG\_LIB=UI-R-EP0  
 TAG\_TISSUE=duodenum  
 TAG\_SEQ=TTGTGTTCAAT\*  
 BASE COUNT 165 a 139 c 219 g 163 t 1 others  
 ORIGIN  
 Query Match 41.8%; Score 488.8; DB 14; Length 687;  
 Best Local Similarity 87.9%; Pred. No. 3.1e-116;

Matches 590; Conservative 0; Mismatches 73; Indels 8; Gaps 5;  
 OY 496 GCGTGGCTGGAAGATTTCTAGCCACAGCAGCAAGCTCTCTGGGCCCGACGCTT 555  
 Db 687 GCGCTGCTGGAAGATTTCTAGCCACAGCAGCAAGCTCTCTGGGCCCGACGCTT 628  
 OY 556 GTGGCAGGTGTCTGGGCGTGTGGCGCTGGGCGAGGGGCTTCCCTTGGATCCGACCT 615  
 Db 627 GTGGCAGGTGTCTGGGCGTGTGGCGCTGGGCGAGGGGCTTCCCTTGGATCCGACCT 568  
 OY 616 CCCCTGGGCTCATCTTAAGGCTGCCCTTCCCTTAACCTTGTGGACTTTCAAGTTCA 675  
 Db 567 CCCCTGGGCTCATCTTAAGGCTGCCCTTCCCTTAACTTGTGGACTTTCAAGTTCA 508  
 OY 676 CTGAGGGGCTTGGCTCTCCAGATTTCTTAACATTTTCC-----TGGCTCAGAGCATCA 731  
 Db 507 CTGAGGGGCTTGGCTCTCCAGATTTCTTAACATTTTCAACAGCTCCGACGACCA 448  
 OY 732 CCACACCTCCCTTACCCCAACCCCACTCCCTCCAGCCCGCTGCTCTTGGTCCAGTCT 791  
 Db 447 CAGACCTCCCAACCCCAACCCCACTCCCTCCAGCCCGCTGCTCTTGGTCCAGTCT 388  
 OY 792 GT-CTCTCTCAAGGACGACGAGCTTGTACATG-TTTCATTCACAGAGCATATC 849  
 Db 387 GTCTCTCTCAAGGACGACGAGCTTGTACATG-TTTCATTCACAGAGCATATTC 328  
 OY 850 TTGCTCTTCT-TAACATCCCATCCACCAACATATCCACTCTACCTCCCAAGGCC 908  
 Db 327 CGGTCTCTCTTAACCTTCCACCCACCAAGCTGCGGCTTCCAGACTCCCAAGGCC 268  
 OY 909 CTACTTATCCCTGACCTCCCAACCCCACTCCACCCAGCTGTATTATGACTTTGTGAC 968  
 Db 267 CCACGTA-CCGAGTACCTCCCAACCCCACTCCAGCTGTATTATGACTTTGTGAT 209  
 OY 969 CAGGACGTAGATGGGCTGACCTGTGTGGCAGGAAGCAGAACTGGGACTAGGCCAG 1028  
 Db 208 CAGGACGTAGATGGGCTGACCTGTGTGGCAGGAAGCAGAACTGGGACTAGGCCAG 149  
 OY 1029 AAGTCCCACTGTGAGGGGGAAGAGCTGGGGAAGAGTCTCCCTGGATCCGAT 1088  
 Db 148 AAGTCCCACTGTGAGGGGGAAGAGCTGGGGAAGAGTCTCCCTGGATCCGAT 89  
 OY 1089 TTTGAAAAGATCTATTTTATTTATTTGACAAAATGTTAATGATTTAAAGAGA 1148  
 Db 88 TTTGAAAAGATCTATTTTATTTATTTGACAAAATGTTAATGATTTAAAGAGA 29  
 OY 1149 ATTAATCTATGA 1159  
 Db 28 ATTAATCTATGA 18  
 RESULT 9  
 LOCUS AW917574 584 bp mRNA linear EST 25-MAY-2000  
 DEFINITION EST348878 Rat gene index, normalized rat, norvegicus, Bento Soares  
 Rattus norvegicus cDNA clone RGIEF49 5' end, mRNA sequence.  
 ACCESSION AW917574  
 VERSION AW917574.1 GI:8083328  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 584)  
 AUTHORS Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (RESTR) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
This clone is available through the ATCC, contact the ATCC, tel#703-365-2700 for further information  
Seq primer: M13 Reverse.

## FEATURES

```

SOURCE
1. .384
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RGIEF4g"
/clone_11b="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/site="Vector: pbluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Estimated insert size approx.1 kb"
BASE COUNT      108 a      188 c      147 g      138 t      3 others
ORIGIN

```

Query Match	41.2%;	Score 481.2;	DB 10;	Length 584;
Best Local Similarity	91.3%;	Pred. No. 2.8e-114;		
Matches 532;	Conservative 0;	Mismatches 46;	Indels 5;	Gaps 2

QY	272	CCTGGCCAGACACAGATGTGAGACAAGCAGGTGTGGATGGAGCACTGATGAGTGGTGGAA	331
Db	2	CCACAGCCAGACAGATGTGACACAGCAGCAGGTGTGGATGTGACACTGATGAGTGGGAA	61
QY	332	GAGACCAAAATCAACAGCTCCAGCCCTCTGCGGTACGACGCGCCAGATTGGGGAAATTTACA	391
Db	62	GAGACCAAAATCAACAGCTCCAGCCCTCTGCGGTATGACGCGCCAGATTGGGGAAATTTACG	121
QY	392	GTCATCAGGGCTGGGGCTCTACTACTCTTACTGTGAGGTGCATCTTTATATAGGAAAGCT	451
Db	122	GTCATCAGGGCTGGGGCTCTACTACTCTTACTGTGACGGTGCATCTTTATATAGGGAAGCA	181
QY	452	GTCACCTGAAGCTGGACTGTGCTGTAAAGGTGTGTGTGCCCCCTGGCGTCCGCGAAGAA	511
Db	182	GTCACCTGAAGCTGGACTGTGCTGTAAATGTGTGTGTGTGCCCCCTGGCGTCCCTGGAGAA	241
QY	512	TTCACAGCACACAGCAGCAGCTCTCCTGCGGCCACCACTCCGTTTGTGCCAGGTGTCTGG	571
Db	242	TTCACAGCACACAGCAGCAGCTCTCCTGCGGCCACCACTCCGTTGTGCCAGGTGTCTGG	301
QY	572	CTGTTGGCCGCTGGGGCCAGGGTCTTCCCTTCGGATCCGACCCCTCCCTGGCTCATCTT	631
Db	302	CTGTTGGCTCTGGGCCAGGGTCTTCCCTTCGGATCCGACCCCTCCCTGGGCTCATCTT	361
QY	632	AAGGCTGCCCCCTCCTAACCTACTTTGGACTCTTTCAAAGTTACTAGAGGGGCGCTTGC	691
Db	362	AAGGCTGCCCCCTCCTTACTTACTTATTTGGACTCTTTCAAAGTTACTAGAGGGGCTCTGC	421
QY	692	TCCAGATTCTTAACTTTCC---TGCGTCCAGAGGATCACCACACTTCCCTCCACC	747
Db	422	TCCAGATTCTTAACTTAACTTAACCTAACACTCCAGAGGATCACCACACTTCCCTCCACC	481
QY	748	CACCCCACTCTCACCCCTGCGTGTCTTGTGTGTCAGTCTGTCTCTCTCAAAAGC	807
Db	482	CACCCCACTCTCACCCCTGCGTGTCTTGTGTGTCAGTCTGTCTCTCTCTCAAGGC	541
QY	808	AGCCAGAGCTTGTCACATG-TTTCATTTCCACAGACGATTC	849
Db	542	AGCCAGAGCTTGATGACATGTTTTCATTTCCACAGACGATATTC	584

RESULT	10
BIR19200	
LOCUS	777 bp mRNA linear EST_04-OCT-2001
DEFINITION	603038361AFL NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5'
ACCESSION	mRNA sequence.
VERSION	BIR19200.1 GI:15930750

**KEYWORDS**  
**SOURCE**

EST.  
human.

**ORGANISM**

Homo S

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi, *Homo*

## REFERENCE

1 (bases 1 to 777)

## AUTHORS

NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE  
JOURNAL.

National Institutes of Health,  
Bethesda, Maryland 20892

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Issue Procurement: Life Technology

CDNA Library Arrayed by: The I.M.A.G.E. Co

DNA Sequencing by: Incyte Genomics, Inc.

clone distribution: MGC clone distribution information can be found through the T M A C E Consortium/INT at:

found image.1111.gov

Plate: LLAM11437 row: 1 column: 03

High quality sequence stop: 759.

LOCATION/QUALITIES  
1-777

/organism="Homo sapiens"

## FEATURES

```

source
1. 777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/clone_lib="NH_MGC_115"
/ab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV.SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NH_MGC Library."
140 a 268 c 215 g 154 t

```

Query Match	41.18;	Score 480.4;	DB 13;	Length 777;
Best Local Similarity	81.68;	Pred. No. 5.1e-114;		
Matches 594;	Conservative	0;	Mismatches 126;	Indels 8;
				Gaps .3

QY	113	GCAGAGGACCGCGGGAGGCCCCCTGGAACATCGAAATCCCCACAGAGAGGAAAGGACAGATGTC	112
Db	14	GCAGAGGAGGACCGAGACCCGTCGGAATCGAATCCCCAGACAGAAAGGACGAGATCT	73
QY	173	GTAACCTTCTTGGAAACAATACTAGTCGGGCTCCAAABAATGCTCTCTMAAGGCCGGAAGCGC	232
Db	74	GCGCGTTCTCTGAACCGCATAGTTGGGCTCCACAAATGGACCTTAAAGGCCGGGAAAAACA	133
QY	233	CGGCGTCGCGGACCTATTGACAGCCATTATGAGTTCATCTCGGGCAGACAGATGGA	292
Db	134	CGGGCTCGAAGACGATCGCAGCCCTTATGAAATTCAATCCACGACCTGGACAGAGACGA	193
QY	293	GCACAAACAGGTGTGATGGAGACAGTGAAGTGGCTGGGAAGAGACCAAAATCAACAGCTCC	352
Db	194	GCGAGGACAGTGTGGACGGAGAGTGAAGTGGCTGGGAGGAAGCAATCAACAGCTCC	253
QY	353	AGCCCTCTGCGCTACGACCCGCGAGATTGGGAAATTTACAGTCATCAGGGGCTGACTAC	412
Db	254	AGCCCTCTGCGCTACCAACCGCGAGATCGGGAGGATTTAATGTCACCCGGGCTGCTTAC	313
QY	413	TACCTGACGTGAGTGTGACATTATATGAGGGAAGGCTGTCTACCTGAACCTGGAATCG	472
Db	314	TACCTGACGTGAGTGTGACATTATATGAGGGAAGGCTGTCTACCTGAACCTGGAATCG	373
QY	473	CTGGTGAACGGTGTGCTGGCCCTCGGCGCTCCCTGGAAGAATTTCTCAGCCACAGCAGCAAGC	532
Db	374	CTGGTGAATGTGTCTGGCCCTCGGCGCTCCCTGGAAGAATTTCTCAGCCACAGCAGCGGCAAGT	433
QY	533	TCTCTGAGGCCCCAGCTCCGTTTGTGTCAGAGTGTCTGGGCTGTGTCGCGCTCGGGCCAGGG	592

BASE COUNT	154 a	287 c	223 g	170 t
ORIGIN				
Query Match	40.5%;	Score 473;	DB 13;	Length 834;
Best Local Similarity	78.0%;	Pred. No. 4.4e-112;		

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 948)  
REFERENCE  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
AUTHORS

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bhs-remail.nih.gov  
Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

FEATURES  
source 1. 948  
Plate: L10M2466 row: n column: 17  
High quality sequence start: 24  
High quality sequence stop: 550.  
Location/Qualifiers

1. 948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5278608"  
/clone\_lib="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pOT7; Site:1; XhoI; Site:2;  
ECORI; cDNA made by oligo-dT priming. Directionally cloned  
into EcorI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Query Match

Best Local Similarity 76.2%; Score 467.8; DB 14; Length 948;

Matches 632; Conservative 0; Mismatches 179; Indels 18; Gaps 4;

129 AGCCCTGACAGTGAATCCCGACAGAGAAAGCCAGATGTGTACCTTTTGGAC 188  
1 ACCCGTCGGAAGTGAATCCCGACAGAGAAAGCCAGATGTGTACCTTTTGGAC 60  
189 AACTAGTCGGGCTCGAAGAGTGTCTTAAAGCGGGAAGGCGGCGCTCGCGAGCTA 248  
61 GACTAGTTGGGCTCGAAGAGTGTCTTAAAGCGGGAAGGCGGCGCTCGCGAGCTA 120  
249 TTGACCCCATTTTGAAGTTTCACTCGGCGAGACAGATGTGTGACCAAGAGTGTG 308  
121 TCGACAGCCCATTTTGAAGTTTCACTCGGCGAGACAGATGTGTGACCAAGAGTGTG 176  
309 ATGGGACAGTGAAGTGTGTGGAGAGACCAAAATCAACAGCTCCAGCCCTCGCTACG 368  
177 ACCGGACAGTGAAGTGTGTGGAGAGACCAAAATCAACAGCTCCAGCCCTCGCTACG 236  
369 ACCGACAGATTTGGGGAATTTACAGTATCAGGCGTGGCTCTACTACTGTACAG 428  
237 ACCGACAGATTTGGGGAATTTACAGTATCAGGCGTGGCTCTACTACTGTACAG 296  
429 TGCACATTTGATGAGGGAAGGCTGTCTACCTGAAGTGAAGTGTGTGAGAGTGTG 488  
297 TGCACATTTGATGAGGGAAGGCTGTCTACCTGAAGTGAAGTGTGTGAGAGTGTG 356  
489 TGGCCCTGCGGCTCGTGAAGATTTCTCAGCCACAGCAGACAGTCTCTCGGCGCCAGC 548  
357 TGGCCCTGCGGCTCGTGAAGATTTCTCAGCCACAGCAGACAGTCTCTCGGCGCCAGC 416  
549 TCGGCTTGGCCAGTGTGTGGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 608  
417 TCGGCTTGGCCAGTGTGTGGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 476  
609 GCACCCCTCGGCTCGTGAAGATTTCTCAGCCACAGCAGACAGTCTCTCGGCGCCAGC 668  
477 GCACCCCTCGGCTCGTGAAGATTTCTCAGCCACAGCAGACAGTCTCTCGGCGCCAGC 536  
669 AAGTCTACGAGGCGGCTGTCTCCAGATTTCTTAACTTTCCTCGGCTCC---AGGA 725

537 AGGTCACGAGGAGGCGGCTGTCTCCCGACAGTCCGACAGGCTGCGGCTCCCTCGAC 596  
726 GCATACACACACTCCCTACCCACCCACCCACCCACCCCTCCGCTGCTGCTGCTGCTG 785  
597 AGCTCTGCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656  
786 AGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838  
657 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712  
839 CAGAGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898  
713 TACAGATTTCCACACTTATTATTTACAAATCCCGCCAGCCAAATTCACCATTAAGTT 772  
899 CCCAAGCCCTACTATATCCCTGACGCTGCGCCACCCACCCACCCACCCACCCACCCAC 947  
773 CCCAATTCCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821

RESULT 13

BI596681 828 bp mRNA linear EST 07-SEP-2001  
603243254F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5285892 5',  
mRNA sequence.

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NIH-MGC http://imgc.nhl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bhs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki  
Toshitoki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10M1722 row: k column: 13  
High quality sequence stop: 776.  
Location/Qualifiers

1. 828  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5285892"  
/clone\_lib="NIH\_MGC\_96"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTT-3',  
size selected for average insert size 2.3 kb and  
normalized to 10<sup>6</sup> copies. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 155 a 278 c 223 g 172 t  
ORIGIN

Query Match

Best Local Similarity 39.1%; Score 456.4; DB 13; Length 828;

Matches 550; Conservative 0; Mismatches 106; Indels 10; Gaps 2;

137 GAATGATATCCCGACAGAGAAAGCAGATGTGTACCTTTTGGACAACTACTAGTC 196

```

|||||
Db 44 GAACGATATCCCAAGAGAGAAAGCCAGATATCTGCTTCTGTAACGCACTACTT 103
Oy 197 CGGCTCGAAGAAATGCTCTCTAAAGCCGGAAGCCGCGCTCGCCAGCTATTGCAACC 256
Db 104 CGGGCTGCAAAATGCACTTAAAGCCGGAAGAAACGCGCTCGAAGAGCATATGCAACC 163
Oy 257 CATATGAGTTCATCCGCGGAGACAGATGAGACAGAGAGAGAGAGAGAGAGAGAGAG 316
Db 164 CATATGAGTTCATCCAGCACTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
Oy 317 GTGAGTGGCTGGAGAGAGACCAAAATCAACAGCTCCAGCCCTGCGCTACAGCCGCGAG 376
Db 224 TTGAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
Oy 377 ATTTGGGAATTTACAGTATACAGAGCTGAGCTCTACTACTACTACTACTACTACTACT 436
Db 284 ATCGGGGATTTATAGTACCCGCGGCTGGCTCTACTACTACTACTACTACTACTACTACT 343
Oy 437 GATGAGGAGAAAGCTGTCTACTGAGCTGAGCTGAGTGAAGGAGTGTGCTGGCCCTG 496
Db 344 GATGAGGAGAAAGCTGTCTACTGAGCTGAGCTGAGTGAAGGAGTGTGCTGGCCCTG 403
Oy 497 CGTGCCTGGAAGATTTCTAGCCACAGACAGCAAGCTCTCTGGGCCCCAGCTCGTTTG 556
Db 404 CGTGCCTGGAAGATTTCTAGCCACAGCTGCGGCAAGTTCCCTCGGGCCCCAGCTCGCTC 463
Oy 557 TGCCAGGTGTCTGGGCTGTCCGCTGCGGCGGAGGCTTTCCCTTGGATTCGCAACCTTC 616
Db 464 TGCCAGGTGTCTGGGCTGTCCGCTGCGGCGGAGGCTTTCCCTTGGATTCGCAACCTTC 523
Oy 617 CCTGGGCTCATTTAAGGCTGGGCGCTTCTTAACCTACTTTGGACTTTTCAAGTTAC 676
Db 524 CCTGGGCTCATTTAAGGCTGGGCGCTTCTTAACCTACTTTGGACTTTTCAAGTTAC 583
Oy 677 TGAGGGGCTTGTCTCCAGATTCCTTAACTTCCCTGGCTCCAGAGCATCACCA 736
Db 584 TGAGGGGCTTGTCTCCAGATTCCTTAACTTCCCTGGCTCCAGAGCATCACCA 636
Oy 737 CCTGCTTACCCACCCCACTCTGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
Db 637 GCTCTTGGGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Oy 794 CTCTCC 799
Db 697 CCTCTC 702

RESULT 14
AA221610 471 bp mRNA linear EST 13-FEB-1997
LOCUS my18d09.r1 Barstead mouse heart MRLRB3 Mus musculus cDNA clone
DEFINITION IMAGE:696209 5', mRNA sequence.
ACCESSION AA221610
VERSION AA221610.1 GI:1840863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musineae; Mus.
1 (bases 1 to 471)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theis, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

```

```

FEATURES
source
location/qualifiers
1..471
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:696209"
/clone_1db="Barstead mouse heart MRLRB3"
/sex="mixed"
/tissue_type="heart"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Organ: heart; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker: Site.1: EcoRI; Site.2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer 15'
TGTTCAGATCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
3'; double-stranded cDNA was ligated to Eco RI adaptors
[CTGATTCGCTAC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT 91 a 141 c 128 g 111 t
ORIGIN
Query Match 37.5%; Score 437.8; DB 9; Length 471;
Best Local Similarity 97.0%; Pred. No. 5.1e-103;
Matches 457; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

Oy 274 TCGGCCAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 333
Db 3 TCGGATCCAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 62
Oy 334 GACCAAAATCAACAGCTCCAGCCCTGCTGCTAGAGCCGAGAGTGGGGAATTTACAGT 393
Db 63 GACCAAAATCAACAGCTCCAGCCCTGCTGCTAGAGCCGAGAGTGGGGAATTTACAGT 122
Oy 394 CATCAGGCTGGGCTCTACTACTGTACTGTACAGTGACATTTGATGAGGAGAGGCTGT 453
Db 123 CATCAGGCTGGGCTCTACTACTGTACTGTACAGTGACATTTGATGAGGAGAGGCTGT 182
Oy 454 CTACCTGAAGCTGACCTGTGTGTGACAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 513
Db 183 CTACCTGAAGCTGACCTGTGTGTGACAGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 242
Oy 514 CTCAGCCACAGCAGCAGCTCCGCGGCGCCAGCTCCGTTTGGCCAGGCTGCTGGGCT 573
Db 243 CTCAGCCACAGCAGCAGCTCCGCGGCGCCAGCTCCGTTTGGCCAGGCTGCTGGGCT 302
Oy 574 GTTGGCGCTGGCGCAGGAGTCTCCCTTGGATCCGACCTCCCTGGGCTCATCTTAA 633
Db 303 GTTGC -GCTGGGAGAGGCTTCCCTTGGATCCGACCTCCCTGGGCTCATCTTAA 360
Oy 634 GGTGCCCCCTTCTTACCTACTTGTGACTTTTCAAGTCACTGAGGAGGCTGCTGCTG 693
Db 361 GGTGCCCCCTTCTTACCTACTTGTGACTTTTCAAGTCACTGAGGAGGCTGCTGCTG 420
Oy 694 CCAGATTCCTTAACTTCCCTGCTCCAGAGAGATCCACACACCTCCCTA 744
Db 421 CCAGATTCCTTAACTTCCCTGCTCCAGAGAGATCCACACACCTCCCTA 471

RESULT 15
B0671259 963 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
DEFINITION 5', mRNA sequence.
ACCESSION B0671259
VERSION B0671259.1 GI:21782093

```







PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
 XX Browning JL, Chicheportliche Y;  
 XX MPI: 1998-145619/13.  
 DR P-PSDB; AAM47524.  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 XX cancer, auto-immune disease and immune responses to tissue grafts  
 PS Claim 2; Pages 45-46; 69pp; English.  
 CC The sequence is that encoding mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 XX  
 SQ Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;  
 Query Match 100.0%; Score 1168; DB 19; Length 1168;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-307;  
 Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GCCCAGCTCCGTTGTCAGAGTGTGCGGTGTGCGGTGCGGCGAGGGCTTCCT 600  
 DB 541 GCCCAGCTCCGTTGTCAGAGTGTGCGGTGTGCGGTGCGGCGAGGGCTTCCT 600  
 QY 601 TCGGATCCGACCCCTCCCTGGGCTCATCTTAAGGTGCCCCCTTCTCACTACTTGG 660  
 DB 601 TCGGATCCGACCCCTCCCTGGGCTCATCTTAAGGTGCCCCCTTCTCACTACTTGG 660  
 QY 661 ACTCTTCAAGTTCATGAGGGGCTTGGCTTCCAGATTCCTTAACCTTCCCTGGCTC 720  
 DB 661 ACTCTTCAAGTTCATGAGGGGCTTGGCTTCCAGATTCCTTAACCTTCCCTGGCTC 720  
 QY 721 CAGGACATACACACACCTCCCTACCCACCCCTCCCTCCACCCCTCCCTCCCTCC 780  
 DB 721 CAGGACATACACACACCTCCCTACCCACCCCTCCCTCCACCCCTCCCTCCCTCC 780  
 QY 781 GGTCCAGTCT 840  
 DB 781 GGTCCAGTCT 840  
 QY 841 GAGTATCTCTGCT 900  
 DB 841 GAGTATCTCTGCT 900  
 QY 901 CAAAGCCCTACTTATCTCTGACTCCCTCCACCCCTCCACCCCTCCCTCCCTCCCT 960  
 DB 901 CAAAGCCCTACTTATCTCTGACTCCCTCCACCCCTCCACCCCTCCCTCCCTCCCT 960  
 QY 961 TTGTGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1020  
 DB 961 TTGTGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1020  
 QY 1021 TAGGCCAGAAAGTTCCTCACTGTGAGGGGAGAGAGTGGGAGAACCTCTCCCTGATCC 1080  
 DB 1021 TAGGCCAGAAAGTTCCTCACTGTGAGGGGAGAGAGTGGGAGAACCTCTCCCTGATCC 1080  
 QY 1081 CTGTGATTTTGAAGAAAGTATTTTATTTATTTATTTATTTATTTATTTATTTAT 1140  
 DB 1141 TTAAGAAATTAATCATGATTTCTCTTC 1168  
 DB 1141 TTAAGAAATTAATCATGATTTCTCTTC 1168  
 RESULT 2  
 AAX23425  
 ID AAX23425 standard; DNA; 701 BP.  
 AC AAX23425;  
 XX  
 DT 18-JUN-1999 (first entry)  
 DE Mouse TNRL3 DNA.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW AP06; AP08; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 OS Mus sp.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..636  
 FT /tag- a  
 FT /product- "TNRL3"  
 XX  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX



PF 04-SEP-1998: 98MO-US18933:  
 XX  
 PR 05-SEP-1997: 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR P-PSDB: AAW93591.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PS  
 PS Example VII: Fig 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TRRL1 and TRRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX  
 SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Query Match 59.9%; Score 699.4; DB 20; Length 701;  
 Best Local Similarity 99.9%; Pred. No. 1.1e-179;  
 Matches 700; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 CTGGTGTGCTACGCTGGGAGCTGGGCAACCTGTCTGCCAGAGCCTTTCAGAG 103  
 DB 1 CTGGTGTGCTACGCTGGGAGCTGGGCAACCTGTCTGCCAGAGCCTTTCAGAG 60  
 QY 104 GACCTGACAGCAGAGGACCGCGGAGCCCGCTGAAGTGAATCCAGACAGAGAAAGC 163  
 DB 61 GACCTGACAGCAGAGGACCGCGGAGCCCGCTGAAGTGAATCCAGACAGAGAAAGC 120  
 QY 164 CAGAGATGTGTACTTCTTGAACAACACTAGTCCGGCTTCGAAGAAGTCTCTAAAGC 223  
 DB 121 CAGAGATGTGTACTTCTTGAACAACACTAGTCCGGCTTCGAAGAAGTCTCTAAAGC 180  
 QY 224 CGGAAGCGCGGCTTCGCGAGCTATTCAGCCCATTTAGAGGTTTATCTCGGACAGA 283  
 DB 181 CGGAAGCGCGGCTTCGCGAGCTATTCAGCCCATTTAGAGGTTTATCTCGGACAGA 240  
 QY 284 CAGAGTGAAGCACAAGAGGTGTGATGGGAGCAGTGAAGTGGGAGAGCAAAATC 343  
 DB 241 CAGAGTGAAGCACAAGAGGTGTGATGGGAGCAGTGAAGTGGGAGAGCAAAATC 300  
 QY 344 AACAGTCCAGCCCTCTGCGCTAGCAGCCGACAGATTGGGAATTTACATCAGTACAGGCT 403  
 DB 301 AACAGTCCAGCCCTCTGCGCTAGCAGCCGACAGATTGGGAATTTACATCAGTACAGGCT 360  
 QY 404 GGGCTCTACTACTGTACTGTCTAGGTGACATTTGATGAGGAAAGGCTGTCTACTGAG 463  
 DB 361 GGGCTCTACTACTGTACTGTCTAGGTGACATTTGATGAGGAAAGGCTGTCTACTGAG 420

QY 464 CTGACTGTGCTGTAAGCGTGTGCTGCGCCCTGCGCTGGAAGAATTCACAGACA 523  
 DB 421 CTGACTGTGCTGTAAGCGTGTGCTGCGCCCTGCGCTGGAAGAATTCACAGACA 480  
 QY 524 GCAGCAGAGCTCTCTGCGGCCCGCAGCTCCGTTTGGCAGAGTGTGGGCTGTGCGCTG 583  
 DB 481 GCAGCAGAGCTCTCTGCGGCCCGCAGCTCCGTTTGGCAGAGTGTGGGCTGTGCGCTG 540  
 QY 584 CGGCCAGAGGCTCTCTGCGGCCCGCAGCTCCGTTTGGCAGAGTGTGGGCTGTGCGCTG 643  
 DB 541 CGGCCAGAGGCTCTCTGCGGCCCGCAGCTCCGTTTGGCAGAGTGTGGGCTGTGCGCTG 600  
 QY 644 TTCTTACCTACTTGTGAGCTTTCAGTTCACTGAGGGGCTGTGCTCCAGATTCT 703  
 DB 601 TTCTTACCTACTTGTGAGCTTTCAGTTCACTGAGGGGCTGTGCTCCAGATTCT 660  
 QY 704 TAACTTCCCTGCTGCAGAGCAGACATCACACATCCCTCA 744  
 DB 661 TAACTTCCCTGCTGCAGAGCAGACATCACACATCCCTCA 701

RESULT 3  
 AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 XX  
 AC AAA49717;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 cDNA clone DNA30879-1152.  
 XX  
 KW PRO207; human; antitumor; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukemia; neoplasm; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide 58..177  
 FT mat\_peptide 178..804  
 FT /\*tag= c  
 PN WO200037638-A2.  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999: 99MO-US28565.  
 XX  
 PR 22-DEC-1998: 98US-0113296.  
 PR 08-MAR-1999: 99MO-US05028.  
 PR 21-APR-1999: 99US-0130232.  
 PR 28-APR-1999: 99US-0131445.  
 PR 14-MAY-1999: 99US-0134287.  
 PR 20-JUL-1999: 99US-0144758.  
 PR 26-JUL-1999: 99US-0145698.  
 PR 15-SEP-1999: 99MO-US21090.  
 PR 15-SEP-1999: 99MO-US21547.  
 XX  
 PA (GENTH ) GENTECH INC.  
 XX  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pittli RM, Wood WI;  
 XX  
 DR WPI: 2000-442668/38.  
 DR P-PSDB: AAY95358.  
 XX  
 PT Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or











PN W09835061-A2.  
 XX 13-AUG-1998.  
 PD 12-FEB-1998; 98WO-US02859.  
 PF 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX (ABBO ) ABBOTT LAB.  
 PA Willey SR;  
 PI WPI: 1998-447255/38.  
 DR P-PSDB; AAM29745.  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 treatment of autoimmune disease, tumours and inflammation  
 PS Claim 11; Page 123-4; 142pp; English.  
 XX The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;  
 SQ  
 Query Match 51.2%; Score 597.8; DB 19; Length 1236;  
 Best Local Similarity 75.2%; Pred. No. 5.6e-152;  
 Matches 879; Conservative 0; Mismatches 222; Indels 68; Gaps 8;

Db 493 TGTGAGTGCACTTTGATGAGGGAGGCTGTACTGAGAGCTGACTTGTGTGGAT 552  
 Qy 482 GGTGTGCTGGGCGCTGGCTGTGAGAAATTCACGACAGACAGAGCTTCCTGGG 541  
 Db 553 GGTGTGCTGGGCGCTGGCTGTGAGAAATTCACGACAGAGCTTCCTGGG 612  
 Qy 542 CCCCAGCTCCCTTTGTCAGAGTGTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCT 601  
 Db 613 CCCCAGCTCCCTTTGTCAGAGTGTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCT 672  
 Qy 602 CGATCCGACCTCCCTGGGCTGATGAGCTGGCTGGCTGGCTGGCTGGCTGGCT 661  
 Db 673 CGATCCGACCTCCCTGGGCTGATGAGCTGGCTGGCTGGCTGGCTGGCTGGCT 732  
 Qy 662 CTCCTTCAAGTTCAGTGGGCGCTGTGCTCCAGATCTCTTAACCTTCCCTGGCT 721  
 Db 733 CTCCTTCAAGTTCAGTGGGCGCTGTGCTCCAGATCTCTTAACCTTCCCTGGCT 792  
 Qy 722 AGAGCATCACCACACCTCCCTGACCCGACCTCCCTGACCCCTGCTGCTGCT 780  
 Db 793 CTTCCGACCTCTCTGTGGGACCGGCTCCCTGCTCCCTGACCCCTGCTGCTGCT 847  
 Qy 781 GGTCCAGTCTGCTCTCTCC--TCAAGGACGACGAGCTGTTCATGTTCCATTC 837  
 Db 848 GCTCCAGACCTGCTCCCTCTGAGAGCTGCTGGGCTGTTCACGTTTCCATTC 907  
 Qy 838 -----ACAGCATCTCTCTCTTAACTCCATCCACCACTATCCACTC 891  
 Db 908 ACATTAATACATGATTTCCACTCTTACTTACCAACCCACGCGCTCCACTC 967  
 Qy 892 ACTAGTCCCAAGCCCTAC-----TTATCCCTGATCCCTCCACCTACT 936  
 Db 968 ACTAGTCCCAAGCCCTAC-----TTATCCCTGATCCCTCCACCTACT 1027  
 Qy 937 CACCGACACGCTGTTATGACTTGTGAC----- 968  
 Db 1028 CAGACCCGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087  
 Qy 969 -----CAGGACGTGAGATGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023  
 Db 1088 CACTTACGACCTAAGAGGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147  
 Qy 1024 GCCAAGTTCACCACTGTGAGGGGAGAGAGCTGGGAGAGCTGCTGCTGCTGCT 1079  
 Db 1148 GCCAAGTTCACCACTGTGAGGGGAGAGAGCTGGGAGAGCTGCTGCTGCTGCT 1206  
 Qy 1080 CCTGTGATTTGAAAGATCTATTTT 1108  
 Db 1207 CCTGTGATTTTAAACAGATATTTT 1235  
 RESULT 9  
 AAD04350  
 ID AAD04350 standard; cDNA; 1236 BP.  
 XX AAD04350;  
 AC 04-JUL-2001 (first entry)  
 DT  
 XX Human TREPA (TNF related endothelium proliferative agent) cDNA.  
 DE Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulnery; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key location/Qualifiers  
 FT CDS 1..750  
 FT /product- "Human TREPA (TNF related endothelium  
 FT proliferative agent)".  
 XX



PT Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PT  
XX  
XX  
PS Example 2; Column 73-74; 53pp; English.  
XX  
XX  
CC The present invention relates to extracellular signal molecules,  
CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TRBPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularizing grafted  
CC tissue for successful grafting and to promote tissue grafts.  
CC The present sequence is a cDNA clone ID #90050 encoding human TREPA.  
XX  
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

[illegible]

QY	542	CCCCGAGTCCGTTTGGCAGAGTGTCTGGGGCTGTGGCGTGCAGGAGGCTTCCCTT	601
Db	613	CCCCAGCTCCGCTTCTCCAGAGTGTCTGGGGCTGTGGCGTGCAGGAGGCTTCCCTT	672
QY	602	CGGATCCGACCCCTCCCTGGGCTCATCTTAAAGGTGCCCCCTTCTTAACTTCTTGA	661
Db	673	CGGATCCGACCCCTCCCTGGGCTCATCTTAAAGGTGCCCCCTTCTTAACTTCTTGA	732
QY	662	CTCTTTCAAGTTCACTGAGAGGGGCTTGTCTCTCCCAAGATTCTTAAACTTCCCTGGCTTC	721
Db	733	CTCTTTCAAGTTCACTGAGAGGGGCTTGTCTCTCCCAAGATTCTTAAACTTCCCTGGCTTC	792
QY	722	AGGACATCACACACATCTCCCTGACCCCAACCCCACTCTCCCAACCCCTC-GETGGCTCTT	780
Db	793	AGGACATCACACACATCTCCCTGACCCCAACCCCACTCTCCCAACCCCTC-GETGGCTCTT	847
QY	781	GGTCAGTCTGTCTCTCTCC--TCAAAGGCAAGCAGAGCTGTGTACATGTTCATTC-837	
Db	848	GGTCAGTCTGTCTCTCTCC--TCAAAGGCAAGCAGAGCTGTGTACATGTTCATTC-837	907
QY	838	-----ACAGACGTATCTCTGTCTCTTCTTAAATCCCATCCACACAACTATCCACCTTC	891
Db	908	ACATTAATACAGTATTTCCCATCTTATCTTACACAAACCCCAACCCCACTCTCCACCTTC	967
QY	892	ACTAGCTCCCAAGGCGCTTAC-----TTATCCCTGACTCCCAACCACT	936
Db	968	ACTAGCTCCCAAGGCGCTTAC-----TTATCCCTGACTCCCAACCACT	1027
QY	937	CACCGACACAGCTGTATTGACTTTGTGCAC-----	968
Db	1028	CAGACCCCAAGGCGCTTGTGTTCACTGACTGTGTGGGCAAGATGGGTCCAGAAAGACC	1087
QY	969	-----CAGGACCTGAGATGGGCTGTGACCTGTGTGGGCAAGATGGGTCCAGAAAGACC	1023
Db	1088	CACCTCAGGACACTAAGAGGGGCTGTGACCTGTGTGGGCAAGATGGGTCCAGAAAGACC	1147
QY	1024	GCCAGAGTTCCCACTGAGAGGGGCAAGAGCTGTGGCAACAGCTCTCCCTGTA-----TC	1079
Db	1148	GCCAGAGTTCCCACTGAGAGGGGCAAGAGCTGTGGCAACAGCTCTCCCTGTA-----TC	1206
QY	1080	CCTGTGATTTTGAAGAAGATACTATTTT 1108	
Db	1207	CCTGTGATTTTGAAGAAGATACTATTTT 1235	
RESULT 10			
AAAX23424			
AC	AAAX23424	standard; DNA; 1030 BP.	
AC	AAAX23424;		
DT	18-JUN-1999	(first entry)	
DE	Human TNRL3 DNA.		
KW	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04; development abnormality; gestational abnormality; prostate cancer; AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease; cytoplasmic domain; immunogen; antibody preparation; breast carcinoma; apoptosis; human; ss.		
OS	Homo sapiens.		
FT	Key	Location/Qualifiers	
FT	CDS	1..627	
FT	/*tag= a		
FT	/product= "TNRL3"		
PN	W09911791-A2.		
XD	11-MAR-1999.		



PF 04-SEP-1998; 98WO-US18393.  
 XX 05-SEP-1997; 97US-0924634.  
 XX (UNIV ) UNIV WASHINGTON.  
 PA Chaudhary PM;  
 PI WPI: 1999-205191/17.  
 DR P-PSDB; AAM93590.  
 XX  
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 XX Example VII; Fig 13A. 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 XX Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:  
 S0  
 Query Match 44.8%; Score 522.8; DB 20; Length 1030;  
 Best Local Similarity 81.6%; Pred. No. 1.2e-131;  
 Matches 643; Conservative 0; Mismatches 137; Indels 8; Gaps 3;  
 OY 53 GCACCCCTGGGAGCTGGGCAACGCTGCTGCCAGAGACCTCTTCAGAGAGAGCTGACA 112  
 DB 1 GTCACCTTGGGGAGCGCGGCATCGCTGCCAGAGACCTGCCAGAGAGAGCTGGTG 60  
 OY 113 GCAGAGAGACCGCGGAGACCCCTGCACTGAATCCCGAGAGAGAAAGCCAGAGATG 172  
 DB 61 GCAGAGAGAGACCGAGACCCCTGCGAATGATCCCGAGAGAGAAAGCCAGATCT 120  
 OY 173 GTCACCTTCTTGGAAACAACCTAGTCCGGCTCGAGAGAGTCTCTAAAGCCCGAAGGCG 232  
 DB 121 GCGCCTTTCCTTAACCCGACTAGTTCGGCTCGCAGAGAGCTCAAAAGCCGGAACAA 180  
 OY 223 GCGGCTCCCGGAGCTATGACAGCCCATATAGTTCATCCCGGCGCAGAGAGATGGA 292  
 DB 181 CGGGCTCGAAGAGCATGCGACGCCCATATGAAATTCAATCCAGACTGACAGAGAGCA 240  
 OY 293 GCACAGAGAGAGTGTGATGAGACAGTGAAGTGGGAGAGAGACAAATCAACAGCTCC 352  
 DB 241 GCGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 OY 353 AGCCCTCTGCGGTACGACCGCAGATGGGGAATTTACAGTATCAGGCGTGGCTCTAC 412  
 DB 301 AGCCCTCTGCGGTACTACGCGAGATCGGAGATTATAGTCAACCGGCTGGCTCTAC 360  
 OY 413 TACCTGTACTGTACAGTGTGATGAGGAGAAAGGCTGTACTGAGTGAAGTCTTG 472  
 DB 361 TACCTGTACTGTACAGTGTGATGAGGAGAAAGGCTGTACTGAGTGAAGTCTTG 420

OY 473 CTGCTGAACGGTGTGCTGGCCCTGGCTGCTGCGAAGAATTCAGCCACAGACAGCAAGC 532  
 DB 421 CTGCTGAATGCTGTGCTGGCCCTGGCTGCTGCGAAGAATTCAGCCACAGTGGCCAGT 480  
 OY 533 TCTCTGGGCCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTGCGCCGCGGCGAAGG 592  
 DB 481 TCCCTCGGGCCCCAGCTCCGCTCTGCGAGAGTGTGGGCTGTGCGCCGCGGCGAAGG 540  
 OY 593 TCTTCCCTGGGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTCTCTAC 652  
 DB 541 TCCCTCCGCGGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTCTCTAC 600  
 OY 653 TACTTGGAGCTTCTCAAGTTCACAGAGAGGCGCTGTCTCCAGATCTCTTAACCTTTC 712  
 DB 601 TACTTGGAGCTTCTCAAGTTCACAGAGAGGCGCTGTCTCCAGATCTCTCTACAGCT 660  
 OY 713 CCTGCTCCAGAGCATACACACCTCCCTACCCACCCCTCCCTCCACACCCCTC-G 771  
 DB 661 GCGGCTCC-----CCTGACAGAGCTCTGCGGACCGCTCCCTCCCTCCACCCCTCAG 715  
 OY 772 CTGCTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 829  
 DB 716 CCGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775  
 OY 830 TCCATTC 837  
 DB 776 TTCCATTC 783  
 RESULT 11  
 ID AAS03964  
 AAS03964 standard; DNA; 898 BP.  
 XX  
 AC AAS03964;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA.  
 XX  
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 fusion protein.  
 OS Homo sapiens.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 52..873  
 FT /tag= a  
 FT /product= "Fusion protein comprising a growth hormone  
 FT leader, a leucine zipper multimerisation  
 FT domain, and human TWEAK extracellular  
 FT domain"  
 XX  
 XX MO200145730-A2.  
 XX  
 XX 28-JUN-2001.  
 XX  
 XX 19-DEC-2000; 2000WO-US34755.  
 XX  
 XX 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 XX (IMMUNEX CORP.  
 XX

PI Wiley SR:

DR WPI: 2001-417975/44.  
P-PSDB: AAU03499.

Modulating angiogenesis in a mammal for treating diseases mediated by angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or peripheral tissue, by administering antagonist or agonist of TWEAK receptor

Example 1; Page 39-40; 46pp; English.

The sequence represents a DNA from the expression vector PDC409-12-TWEAK, which encodes a fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and the extracellular domain of human TWEAK. The fusion protein was used in the isolation of human TWEAK receptor (TWEAKR)-expressing clones from a COS cell human cDNA library. The TWEAK protein is a member of the tumour necrosis factor (TNF) family and induces angiogenesis. TWEAKR may therefore be used to screen for and develop TWEAKR agonists and antagonists for the modulation of angiogenesis, to be used in the treatment and diagnosis of human disease. The disorders mediated by angiogenesis include ocular disorders characterised by ocular neovascularisation such as diabetic retinopathy, neovascular glaucoma, retinoblastoma, retinopathy of prematurity, retrolental fibroplasia, rubeosis, uveitis, macular degeneration and corneal graft neovascularisation, and inflammatory diseases such as arthritis, rheumatism and psoriasis. Other treatable diseases include malignant and metastatic conditions such as sarcomas and carcinomas, benign tumours and preneoplastic conditions, myocardial angiogenesis, haemophilic joints, scleroderma, vascular adhesions, atherosclerotic plaque neovascularisation, telangiectasia, wound granulation, coronary atherosclerosis, peripheral atherosclerosis and ischemia.

Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:

Query Match 42.7%; Score 498.8; DB 22; Length 898;  
Best Local Similarity 87.0%; Pred. No. 3.7e-125;  
Matches 548; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 56 AGCCGGGAGAGTGGGACACGCTGTGCTCCAGAGAGCTTTCAGAGAGCTGACACCA 115  
DB 250 AGTTGGGGAGCCGGGCTGCTGTCCCGCAGAGAGCTTCCAGAGAGCTGTTGGCA 309  
QY 116 GAGGACCCGCGGAGCCCTGAACTGAATCCAGAGAGGAAAGCCAGAGTGTGTA 175  
DB 310 GAGGAGGACAGAGACCCCTCGAACTGAATCCAGAGAGGAAAGCCAGAGTCTCG 369  
QY 176 CTTTCTTGAACAACAGTACGCTCGGCTTCGAAGAGCTCTAAAGCCGGAAGCGCGG 235  
DB 370 CTTTCTTGAACAACAGTACGCTCGGCTTCGAAGAGCTCTAAAGCCGGAAGAGCGG 429  
QY 236 CCTGGCCAGATATTCAGAGCCATTAATGAGTTCATCTCGGCGAGAGAGATGAGCA 295  
DB 430 GCTGGAAGAGGATTCGAGCCATTAATGAGTTCATCTCGGCGAGAGAGCGG 489  
QY 296 CAACAGAGTGTGATGAGAGAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 355  
DB 490 CAGGAGAGTGTGAG 549  
QY 356 CCTGTGCGCTAG 415  
DB 550 CCTGTGCGCTAG 609  
QY 416 CTGACAGTGTGAG 475  
DB 610 CTGACAGTGTGAG 669  
QY 476 GTGAAGAGTGTGAG 535  
DB 670 GTGAGATGTGTGAG 729  
QY 536 CTTGGGCCCCAGGCTCCGTTTGTGCGAGAGTGTGAGGCTGTGCGGCTGCGGCGAGGCTCT 595

DB 730 CTCGGGCCCCAGGCTCCCTGCTGCGAGAGTGTGAGGCTGTGAGGCTGAGGCGCAGAGGTC 789  
QY 596 TCCCTTCGAGATCGGACACCTCCCGGCGTCAATCTTAAGGCGCGCCCTCTACACATAC 655  
DB 790 TCCCTTCGAGATCGGACACCTCCCGGCGTCAATCTTAAGGCGCGCCCTCTCTACACATAC 849  
QY 656 TTTGGAGCTCTTCAAGTTCACACTGAGAGGCGC 685  
DB 850 TTCGAGCTCTTCAGAGTTCACACTGAGAGGCGC 879

RESUTF 12

AA722190  
ID AA722190 standard; cDNA to mRNA; 282 BP.

AA722190;

27-AUG-1996 (first entry)

Human gene signature HUMG503761.

Gene signature; messenger RNA; mRNA; relative abundance; frequency;

human; cloning; mapping; non-biased library; diagnosis; detection;

cell typing; abnormal cell function; ss.

Homo sapiens.

W09514772-A1.

01-JUN-1995.

11-NOV-1994; 94WO-JP01916.

12-NOV-1993; 93JP-0355504.

(MATS/) MATSUBARA K.

(OKUBO/) OKUBO K.

Matsubara K, Okubo K;

WPI: 1995-206931/27.

Claim 1; Page 1067; 2245pp; Japanese.

A single-stranded DNA (or its complementary strand or the corresp. double-stranded DNA) which comprises one of the 7837 "GS" sequences given in AA719001-726837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.

Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other:

Query Match 7.1%; Score 82.8; DB 16; Length 282;  
Best Local Similarity 80.6%; Pred. No. 1.9e-12;  
Matches 158; Conservative 0; Mismatches 26; Indels 12; Gaps 5;

QY 969 CAGGACACTGATGGGCTGAGCTG-GTGGCAGAGAGCCAGAGACCTGGAGCTAGGCCA 1027

```

Db      88 CAGGACTAAGAGGGGCTGGACCTNTGGCGAGGAGCAAGCAAGAGCTGGCGCTAGGGCA 147
      1028 GAAGTCCCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGCA----TCCTCG 1083
      148 GGAGTCCCAAAATGTGAGGGGCGAGA-AACAAGACAGAGCTCTCCCTGCAAAATCCCTG 206
      1084 TGGATTTTGAAGAA--AGATACATATTTTATATATATGTCGACAAATGTT----AAATGGA 1137
      207 TGGATTTTGAACAGATATATTTTNTATATGTGACAAATGTTGTAATATGGA 266
      1138 TATTAAGAGAAATAAA 1153
      267 TATTAATATAGATAAA 282

RESULT 13
ABK29540
ID      ABK29540 standard; cDNA; 195 BP.
XX
AC      ABK29540;
XX
DT      23-APR-2002 (first entry)
XX
DE      Colon adenocarcinoma-specific cDNA #66.
XX
KW      Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO200196389-A2.
XX
PD      20-DEC-2001.
XX
PF      07-JUN-2001; 2001WO-US18574.
XX
PR      09-JUN-2000; 2000US-210667P.
XX
PR      22-NOV-2000; 2000US-252614P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Meagher MJ, King GE, Xu J, Secrist H;
XX
DR      WPI; 2002-098052/13.
XX
PT      New isolated polynucleotide encoding a polypeptide comprising a portion
PT      of colon tumour protein, for detection, diagnosis and therapy of human
PT      colon cancer -
XX
PS      Claim 1; Page 133; 211pp; English.
XX
CC      The invention relates to an isolated polynucleotide (I) encoding a
CC      polypeptide (II) comprising a portion of a colon tumour protein. A new
CC      oligonucleotide (III) that hybridises to (I) is useful for
CC      determining the presence of a cancer in a patient. (II) or antigen
CC      presenting cells expressing (I) are useful for stimulating and/or
CC      expanding T cells specific for a tumour protein, by contacting T cells
CC      with (I), (II) or antigen-presenting cells that express (I), (I), (II),
CC      or antigen presenting cells that express (II) are useful for treating
CC      colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated
CC      from a patient with (I), (II), or antigen presenting cells that express
CC      (II), so that T cells proliferate, and administering to the patient an
CC      effective amount of the proliferated T cells, thus inhibiting the
CC      development of a cancer in the patient. A new composition is useful for
CC      stimulating an immune response in a patient. (I) or (II) is useful in
CC      vaccines and pharmaceutical compositions for prevention and treatment of
CC      colon cancer and for the diagnosis and monitoring of the cancers. (I),
CC      (II) or an antibody against (II) is useful for detection, diagnosis and/
CC      or therapy of human colon cancer. (I) is useful as a probe or primer for
CC      nucleic acid hybridisation, and in the design and preparation of ribozyme
CC      molecules for inhibiting expression of (II) in tumour cells. ABK29475-
CC      ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of
CC      the invention.
```

```

XX
SQ      Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;
XX
Query Match      6.0%; Score 69.8; DB 24; Length 195;
Best Local Similarity 83.5%; Pred. No. 5,4e-09;
Matches 91; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
XX
      969 CAGGACTAAGAGGGGCTGGACCTGTGGCGAGGAAGCCAGAGACCTGGGACTAGGGCAG 1028
      69 CAGGACTAAGAGGGGCTGGACCTGTGGCGAGGAAGCCAGAGACCTGGGCTAGGGCAG-128
      1029 AAGTCCCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGCA 1077
      129 GAGTCCCAAAATGTGAGGGGCGAGA-AACAAGACAGAGCTCTCCCTTGA 176

RESULT 14
ABN55975
ID      ABN55975 standard; DNA; 65 BP.
XX
AC      ABN55975;
XX
DT      15-JUL-2002 (first entry)
XX
DE      Mouse spliced transcript detection oligonucleotide SEQ ID NO:28723.
XX
KW      Human; mouse; rat; splice transcript; detection; RNA transcript;
KW      splice variant; transcriptome; oligonucleotide library; ss.
XX
OS      Mus musculus.
XX
PN      WO200210449-A2.
XX
PD      07-FEB-2002.
XX
PF      20-JUL-2001; 2001WO-IB01903.
XX
PR      28-JUL-2000; 2000US-221607P.
XX
PR      02-MAY-2001; 2001US-287724P.
XX
PA      (COMP-) COMPUGEN INC.
XX
PI      Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR      WPI; 2002-257383/30.
XX
PT      New oligonucleotide libraries comprising oligonucleotides which
PT      selectively hybridize to mRNAs transcribed from a transcription unit of
PT      a genome, useful for detecting tissue-, pathology-, and
PT      developmental-specific genes -
XX
PS      Example 1; SEQ ID 28723; 47pp; English.
XX
CC      The present invention describes oligonucleotide libraries for detecting
CC      messenger RNAs that populate a (sub-)transcriptome, where the
CC      (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC      transcription units that populate a genome. The library comprises
CC      several oligonucleotides, each capable of hybridising selectively to a
CC      set of messenger RNAs transcribed from a given transcription unit of
CC      the genome, which encodes one or more messenger RNA splice variants.
CC      The oligonucleotide libraries are useful for detecting mRNAs from a
CC      biological sample, in expression profiling studies, in qualitatively or
CC      quantitatively characterising the corresponding transcriptome, and in
CC      detecting RNA transcripts and splice variants of human or animal
CC      transcriptomes. The libraries may also be used as specialised mini
CC      libraries to detect transcripts of a sub-transcriptome under a
CC      particular biological or pathological state, and so allowing the
CC      detection of tissue- and pathology-specific genes such as those genes
CC      only expressed in specific tissue under a specific pathological
CC      condition; to detect developmental specific genes; and to detect RNA
CC      transcripts and splice variants of a transcriptome of a patient suffering
CC      from a particular disorder. ABN27253 to ABN59589 represent
CC      oligonucleotide sequences from rats, humans and mice, which are used in
```

CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 65 BP; 14 A; 15 C; 21 G; 15 T; 0 other;  
Query Match 5.6%; Score 65; DB 24; Length 65;  
Best Local Similarity 100.0%; Pred. No. 6.6e-08;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1030 AGTTCCCACTGTGAGGGGGAAGAGCTGGGACAGAGCTCCTCCCTGGATTCCTGTGATT 1089  
DB 1 AGTTCCCACTGTGAGGGGGAAGAGCTGGGACAGAGCTCCTCCCTGGATTCCTGTGATT 60  
QY 1090 TTGAA 1094  
DB 61 TTGAA 65  
RESULT 15  
AAD30228  
ID AAD30228 standard; DNA; 53522 BP.  
XX  
AC AAD30228;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human PKD1 gene.  
XX  
KW Human; PKD1 gene; autosomal dominant polycystic kidney disease; ADPKD;  
KW acquired cystic disease; transgenic animal; chromosome 16; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200206529-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 13-JUL-2001; 2001WO-US22035.  
XX  
PR 13-JUL-2000; 2000US-218261P.  
XX  
PR 13-APR-2001; 2001US-283691P.  
XX  
PA (UY30 ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
XX  
PI Germlino GG, Watnick TJ, Phakdeekitchareon B;  
XX  
DR WPI; 2002-179805/23.  
XX  
PT Novel primer for diagnosing polycystic kidney disease-associated  
PT disorder, comprises regions having sequence that selectively hybridizes  
PT to polycystic kidney disease gene sequence -  
XX  
PS Claim 20; Page 127-156; 192pp; English.  
XX  
XX The present invention relates to compositions and methods useful for the  
CC identification and detection of polycystic kidney disease (PKD1) gene  
CC mutations. The invention also relates to primers comprising a 5' region  
CC having a sequence that selectively hybridizes to a PKD1 gene sequence  
CC and optionally, to a PKD1 homologue sequence and an adjacent 3' region  
CC having a sequence that selectively hybridizes to a PKD1 gene sequence  
CC and not to a PKD1 homologue sequence. Primer pairs of the invention are  
CC useful for detecting the presence or absence of a mutation in a PKD1  
CC polynucleotide in a sample, for identifying a subject at risk for a  
CC PKD1-associated disorder such as autosomal dominant polycystic kidney  
CC disease (ADPKD) or acquired cystic disease and for diagnosing a PKD1-  
CC associated disorder in a subject. They are useful for selectively  
CC amplifying a region of a PKD1 gene. PKD1 DNA fragments are useful  
CC detecting the presence of a mutant PKD1 polynucleotide in a sample,  
CC as a probe for an amplification reaction. In hybridisation or  
CC amplification assays of biological samples to detect abnormalities  
CC of PKD1 expression and for engineering transgenic animals. The present

CC sequence is human PKD1 gene located on chromosome 16.  
XX  
SQ Sequence 53522 BP; 8486 A; 17664 C; 15766 G; 11606 T; 0 other;  
Query Match 4.0%; Score 46.6; DB 24; Length 53522;  
Best Local Similarity 48.7%; Pred. No. 0.15;  
Matches 127; Conservative 0; Mismatches 134; Indels 0; Gaps 0;  
QY 685 CTGCTCTCTCCAGATTCTTAACCTTCCCTGGCTCCAGAGCATCAGACACTCCCTA 744  
DB 34365 CTTCCCTCTCCCTTACCCCTTTCCCTCCCTCCCTAGACCTTCCCTCAGCTCTCC 34424  
QY 745 CCCCACCCCACTCTCCACCCCTCCCTGCTCTTGTCCAGTCTGTCCTCAAA 804  
DB 34425 CGCTGAGCCCTCCACTCGTCCGAGCCCTCCCTCCCTAGCCCTCCCTCCCTT 34484  
QY 805 GGCAGCAGAGCTTGTTCATGTTCATTCACAGAGATATCTTGTCTTAAACA 864  
DB 34485 CCTCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCTCTCCCTCCCTCCCTCC 34544  
QY 865 TCCCATCCACACAGACTATACACTACACTAGCTCCCAAGCCCTACTTATCCCTGACT 924  
DB 34545 TTCTTCCCTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 34604  
QY 925 CCCCACCCCACTCAGCCGACC 945  
DB 34605 TCCTCCCTCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34625

Search completed: May 8, 2003, 07:30:14  
Job time : 314.523 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:00 ; Search time 2992.86 Seconds

(without alignments)  
11357.738 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtgctgagcctggcctg.....ataatcatgtctcttc 1168

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_din:\*  
35: em\_hlg\_rod:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vrt:\*  
38: em\_sy:\*  
39: em\_hlg\_hum:\*  
40: em\_hlg\_mus:\*  
41: em\_hlggo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1168	100.0	1168	10 AF030100	AF030100 Mus muscu
2	711	60.9	203083	2 AC069459	AC069459 Mus muscu
3	711	60.9	234182	10 AL603707	AL603707 Mouse DNA
4	628.6	53.8	1353	6 AX201324	AX201324 Sequence
5	628.6	53.8	1368	9 AF055872	AF055872 Homo sapi
6	624	53.4	1306	6 AF030099	AF030099 Homo sapi
7	597.8	51.2	1236	6 AR140407	AR140407 Sequence
8	501	42.9	138792	6 AC119115	AC119115 Rattus no
9	498.8	42.7	898	6 AX180714	AX180714 Sequence
10	409.4	35.1	1651	9 BC019047	BC019047 Homo sapi
11	304	26.0	215795	2 AC127470	AC127470 Pan trogl
12	303.8	26.0	177703	2 AC016876	AC016876 Homo sapi
13	278.2	23.8	153553	2 AC126921	AC126921 Bos tauru
14	237.4	20.3	177555	2 AC130192	AC130192 Sus scrof
15	212	18.2	161428	2 AC126925	AC126925 Canis fam
16	156.8	13.4	184026	2 AC098923	AC098923 Rattus no
17	111.6	9.6	203281	2 AC126237	AC126237 Canis fam
18	88.4	7.6	7218	6 I66494	I66494 Sequence 14
19	86.4	7.4	139405	2 AC126239	AC126239 Fells cat
20	69.8	6.0	195	6 AX379024	AX379024 Sequence
21	61	5.2	190376	2 AC123372	AC123372 Rattus no
22	58.4	5.0	64424	2 AC124103	AC124103 Mus muscu
23	56.2	4.8	114260	2 AC098956	AC098956 Rattus no
24	56.2	4.8	125020	9 AF429315	AF429315 Homo sapi
25	55.8	4.8	80112	2 AC128800	AC128800 Rattus no
26	55.8	4.8	303091	2 AC084799	AC084799 Mus muscu
27	55.4	4.7	69119	2 AC128983	AC128983 Rattus no
28	55.4	4.7	241570	2 AC094065	AC094065 Rattus no
29	55	4.7	185822	2 AC073554	AC073554 Homo sapi
30	54.8	4.7	100511	2 AC010774	AC010774 Homo sapi
31	54.6	4.7	221758	2 AC068947	AC068947 Mus muscu
32	54.4	4.7	87120	2 AC012225	AC012225 Homo sapi
33	54.2	4.6	298166	2 AC087563	AC087563 Homo sapi
34	54	4.6	111627	2 AC108626	AC108626 Rattus no
35	54	4.6	184036	2 AC126109	AC126109 Rattus no
36	54	4.6	310371	2 AC096296	AC096296 Rattus no
37	53.8	4.6	99517	2 AC131402	AC131402 Rattus no
38	53.6	4.6	205350	2 AC078946	AC078946 Mus muscu
39	53.4	4.6	146986	2 AC095995	AC095995 Rattus no
40	53.4	4.6	203281	2 AC126237	AC126237 Canis fam
41	53.2	4.6	49430	2 AC100434	AC100434 Mus muscu
42	53.2	4.6	182269	2 AC107416	AC107416 Rattus no
43	53	4.5	55061	2 AC091597	AC091597 Mus muscu
44	53	4.5	176645	2 AC114115	AC114115 Rattus no
45	53	4.5	187252	10 AL607109	AL607109 Mouse DNA

#### ALIGNMENTS

RESULT 1	AF030100	1168 bp	linear	ROD 20-DEC-1997
LOCUS	AF030100	Mus musculus	TWEAK mRNA, partial cds.	
DEFINITION	AF030100			
ACCESSION	AF030100			
VERSION	AF030100.1	GI:2707220		
KEYWORDS				
SOURCE	Mus musculus			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 1168)			
AUTHORS	Chicheportriche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L.			
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family			

Pred. No. is the number of results predicted by chance to have a



## JOURNAL

Submitted (31-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 25, 2001 this sequence version replaced g1:12621364.

## COMMENT

Genome Center

Center: Baylor College of Medicine

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: RP23-168P5

Center clone name: RP23-168P5

Summary Statistics

Sequencing vector: M13: L08821

Chemistry: Dye-terminator Big Dye: 528 of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 212648 bases at least Q40

Consensus quality: 218902 bases at least Q30

Estimated insert size: 222384 bases at least Q20

Quality coverage: 0x in Q20 bases; 249056-tp estimation

Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 7 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 62152: contig of 62152 bp in length

62153 62252: gap of unknown length

62253 118772: contig of 56520 bp in length

118773 148924: gap of unknown length

148925 149024: gap of unknown length

149025 167231: contig of 18207 bp in length

167232 167331: gap of unknown length

167332 189907: contig of 22576 bp in length

189908 190007: gap of unknown length

190008 196537: contig of 6530 bp in length

196538 196637: gap of unknown length

196638 203083: contig of 6446 bp in length.

Location/Qualifiers:

1. 203083

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="11"

/clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others

ORIGIN

Query Match 60.9% Score 711. DB 2: Length 203083;

Best Local Similarity 99.6% Pred. No. 2.5e-176;

Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

425 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 484

DB 43019 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 42960

DB 485 GTGCTGAGCCCTGCGCTGGAAGAATTCAGCCACAGCAGCAAGCTCTCTGGGCC 544

DB 42959 GTGCTGAGCCCTGCGCTGGAAGAATTCAGCCACAGCAGCAAGCTCTCTGGGCC 42900

DB 545 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 604

DB 42899 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 42840

DB 605 ATCCGACACCTCCCTGAGCTCATTTAAGGCTGCCCCCTTCTAACCTATTGACATC 664

DB 42839 ATCCGACACCTCCCTGAGCTCATTTAAGGCTGCCCCCTTCTAACCTATTGACATC 42780

DB 665 TTTCAGATTCACAGAGGGGCTGCTCTCCAGATTCCTTAACTTCTCCGCTCCAGG 724

DB 42779 TTTCAGATTCACAGAGGGGCTGCTCTCCAGATTCCTTAACTTCTCCGCTCCAGG 42720

DB 725 AGATATCACACACTCTCTACCCACCCACATCTCTACACCCCTGCTCTCTGCTC 784

DB 42719 AGATATCACACACTCTCTACCCACCCACATCTCTACACCCCTGCTCTCTGCTC 42660

DB 785 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 842

DB 42659 CAGTGCACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGAGTGAACGGT 42600

DB 843 CGATCTCTGCTCTCTCTTAACTATCCATCCACACCACTATCCACCTGATCC 901

DB 42599 CGATCTCTGCTCTCTCTTAACTATCCATCCACACCACTATCCACCTGATCC 42540

DB 902 AAGCCCTTCTATCT 961

DB 42539 AAGCCCTTCTATCT 42480

DB 962 TGTGACACGAGCAGTGTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1021

DB 42479 TGTGACACGAGCAGTGTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 42420

DB 1022 AGCCAGCAAGTCTCCAACTGTGAGGGGAGAGAGCTGGGAGAGCTCTCTCTCTCTCT 1081

DB 42419 AGCCAGCAAGTCTCCAACTGTGAGGGGAGAGAGCTGGGAGAGCTCTCTCTCTCTCT 42360

DB 1082 TGTGATTTTGAAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1141

DB 42359 TGTGATTTTGAAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 42300

DB 1142 AAGAGATTAATCATGATTTCT 1168

DB 42299 AAGAGATTAATCATGATTTCT 42273

RESULT 3

AL603707/

LOCUS 234182 bp DNA linear ROD 17-NOV-2001

DEFINITION complete sequence.

ACCESSION AL603707

VERSION AL603707.5 GI:17017790

KEYWORDS HTG.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 234182)

REFERENCE

AUTHORS Pearce, A.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Nov 20, 2001 this sequence version replaced g1:16605765.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submissions

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given







D	b670	CCCCAGCTCCGCGTTCGGCAGAGTGTCGGGGCTGTGGGCCATGGGCCCAAGGGCTCCCGCTG	729
O	y	CGGATCCGCACCTCCCTGGGCTCATTTAAGCTGGCCCCCTTCCTTAACCTACTTTGA	661
D	b730	CGGATCCCACCTCCCTGGGCGCCATCTCAAAGCTGGCCCCCTTCCTCACTACTCGGA	789
O	y	CTCTTCAAGTTCACTGAGGGGCGTGTCTGCCAGATTCCCTAACTTTCCTGGGCTC	721
D	b790	CTCTTCCAGGTTCACTGAGGGGCGCTGTCTGCCAGATTCCCTCACTACTCGGCTCC	849
O	y	AGGAGCATCACACACTCTCCCTACCCCAACCCCACTCTCCACCCCTC-CCTGCTCCTT	780
D	b850	-----CCTCGAAGCGCTCTGTGGGACACGGGTCCTCTGCCCCACCTCAGCCGCTCTT	904
O	y	GTCTCAGTCTGCTCTCTCC--TCAAAGCAGACAGAGCTTGTTCAGATGTTTCCATTC-	837
D	b905	GCTCCAGACCTGCCCCCTCCCTCTAAGAGCGTCTGGGCGCTGTTCACGTTTTCCATCC	964
O	y	838 -----AAGAAGTATTCCTTGTCTCTTAACATCCCATCCACACAATTCACCTC	891
D	b965	ACATAAATACGATATTTCCCACTCTTATCTTACACTCCCAACGCGCCACCTTCACCTC	1024
O	y	892 ACTAGCTCCCAAGCCCTTAC-----TTATCCCTATCTCCCAACCCACT	936
D	b1025	ACTAGCTCCCAACCTCCCTGACCCTTTGAGGCCGCCAGTAGTCTGCACTCCCGCTGGCA	1084
O	y	937 CACCGACACGATGTTATGACTTTGTGCAC-----	968
D	b1085	CAGACCCCGAGTCTTGTCTTCACTGTACTCTGTGGCCAAGATGGGTCCAGAACCC	1144
O	y	969 -----CAGGCACTGATGAGGCTGACCTGGGCGGAGGAAGCANAGAACCTGGGACTAG	1023
D	b1145	CACCTCAGGCACTAAGAGGGCTGACCTGGGCGGAGGAAGCAAGAGACTGGGCTTAG	1204
O	y	1024 GCCAAGACTTCCCACTGTGAGGGGAGAGCTGGGGCAAGCTCTCCCTGGA----TC	1079
D	b1205	GCGAGGAGTCTCCCAATGTGAGGGCGAGA-AACAAGCAAGCTCTCTCCCTGGAATTC	1263
O	y	1080 CCTGTGCAATTTTGA--AGATACTATTTTATTTATTTATGTGACAAAATGT--TAAAT	1134
D	b1264	CCTGTGCAATTTTGAAGATATTTATTTATTTATTTATTTGACAAAATGTTGATTAAT	1323
O	y	1135 GGATTTAAGAGATAATCATGA	1159
D	b1324	GGATTTAATGATAATGATCAATA	1348
R	E	S	
L	A	F055872	
D	E	AF055872	1368 bp mRNA linear PRI 04-MAY-1998
I	N	Homo sapiens Apo3/DR3 ligand (Apo3L) mRNA, complete cds.	
V	E	AF055872.1 GI:3108230	
K	E		
S	O	Homo sapiens.	
O	R	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
M	I	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
R	E	1 (bases 1 to 1368)	
A	T	Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and	
T	I	Ashkenazi,A.	
T	I	Identification of a ligand for the death-domain-containing receptor	
J	O	Apo3	
R	N	Curr. Biol. 8 (9), 525-528 (1998)	
M	E	9828355	
P	B	950343	
R	E	2 (bases 1 to 1368)	
A	T	Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and	
A	T	Ashkenazi,A.	
T	I	Direct Submission	
J	O	Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,	
R	E	South San Francisco, CA 94080, USA	
E	A	Location/Qualifiers	

[illegible]



Oy	838	--	---ACAGAGAGTATCCGTTGCTGCTCTCTTTAAATCCCAATCCCAACGACAACTATTCACATCCG	891
Db	925	ACATATAATACGATATATTTCCACTCTTATCTTACAACTCCCCACAGCCCACTCTCCACATCC	984	
Oy	892	ACTAGCTCCCAAGGCCCTTAC-----TTATCCCTACTCTCCCAACCCACT	936	
Db	985	ACTACTCTCCCAATCCCTGACCCCTTTGAGAGCCCCCAGTATCTCGACTCTCCCTCGGCCA	1044	
Oy	937	CACCCGACCAGCTGTTTATTTGACTTTGTGCAC-----	968	
Db	1045	CAGACCCCCAGGGCATTTGTTGTTCACTGTACTCTGTGGCCAGATAGGGTCCAGAAGACCC	1104	
Oy	969	-----CAGGCACTGAGATGGGCTGGACCTGGTGGGACAGGAAGCCAGAAACCTGGGACTAG	1023	
Db	1105	CACTTTCAGGCACTAAGAGGGGCTGGAACCTGGGGGAGGAAGCCAAAGAACATGGGGCTTGG	1164	
Oy	1024	GCCAGAGTGTCCCACTGTGAGGGGGAAGAGCTGGGGACAACCTCCTCCCTGSA-----TC	1079	
Db	1165	GCCAGAGAGTGTCCCAATGTGAGGGGCGAGA-AACAGACAAACCTCCTCCTTGGACGATTC	1223	
Oy	1080	CCTGTGATTTTGA---AGATACATTTTATTTATTTATTTGACAAAATGT---TAAAT	1134	
Db	1224	CCTGTGATTTTATAACAGATATTATTTTTATTTATTTATGTGACAAAATGTGTGATTAAT	1283	
Oy	1135	GGATTTTAAAGGAAATTAATCA	1156	
Db	1284	GGATTTTAAATGAATTAAGTCA	1305	

RESULT 7			
ARI40407	1336 bp	DNA	linear
LOCUS			
DEFINITION	Sequence 1 from patent US 6207642.		PAT 16-JUN-2001

SOURCE	ORGANISM
unknown.	unknown.
unknown.	unknown.
unclassified.	unclassified.
REFERENCE	1 (bases 1 to 1236)

BASE COUNT	ORIGIN	FEATURES	JOURNAL	TITLE
225 a	416 c	358 g	237 t	Member of the TNF family useful for treatment and diagnosis of disease
		1. .1236	US 6207642-A 1 27-MAR-2001;	Patent:
			Location/Qualifiers	
			"/organism="unknown"	

Query Match	51.2%	Score	597.8	DB	6	Length	1236
Best Local Similarity	75.2%	Pred. No. 1e-146					
Matches 879; Conservative	0	Mismatches	222			Indels	68
						Gaps	8

[illegible]

OY		302	GGTGTGATGGGACAGTAGTGGCTGGGAAGAAGACCMAAATTCAAAGCTCCAGCCCTCTG	361
Db		373	TTTTTTTGGAGCGGACAGTAGTGCTGGGAGGAAGCAGATCAAGACTCCAGCCCTTG	432
OY		362	CCTACGACCCGCACATTGGGGAATTTACAGCATCACAGGCTGGGCTCTACTACCTGTAC	421
Db		433	CCCTACCAACCCGCANATCGGGAAGTTTATAGTACCCCGGCTGGGGCTCTACTACCTGTAC	492
OY		422	TCTCAGGTGCATTTGATGATGAGGAAAAGGCTGTCTACTGTAAAGCTGACTTGTGTGAAC	481
Db		493	TCTCAGGTGCATTTGATGATGAGGGAAGGCTGTCTACTGTAAAGCTGACTTGTGTGAT	552
OY		482	GSTGTGCGGCCCTGCGCTGCTGGAAAGATTCACGACACAGCAAGCAAGCTCTCTGGG	541
Db		553	GSTGTGCGGCCCTGCGCTGCTGGAAAGATTCACGACACAGCAAGCAAGCTCTCTGGG	612
OY		542	CCCCAGCTCCCTTTGTGCGCAGGTGTCTGTGGGCTGTTCGCGCTGCGGCACAAGGCTTCCCTT	601
Db		613	CCCCAGCTCCCTTTGTGCGCAGGTGTCTGTGGGCTGTTCGCGCTGCGGCACAAGGCTTCCCTG	672
OY		602	CGGATCCGACCCCTCCCTGGGCTCATCTTAAGCTGCCCTTCCTTAACCTACTTTGGA	661
Db		673	CGGATCCGACCCCTCCCTGGGCTCATCTTAAGCTGCCCTTCCTTAACCTACTTTGGA	732
OY		662	CTCTTTCAGTTCACTGAGGGGCTTGGCTCCGAGATTCTTAACCTTTCCCTGGCTCC	721
Db		733	CTCTTTCAGTTCACTGAGGGGCTTGGCTCCGAGATTCTTAACCTTTCCCTGGCTCC	792
OY		722	AGGAGCATCACACACTCCCTTACCACCCCACTCCCTCCACCCCCTC-GCTGCTCCTT	780
Db		793	-----CCTGACAGCTCTGTGGGCAACCGGCTCCCTTGCGCCACCTCGAGCGGCTCTTT	847
OY		781	GGTCCAATCTGTCTCTCTC--TCAAAGCAGGCAAGACTTGTTCACATTTTCATTGC-	837
Db		848	GCTCCAGACTGCCCCCTCCCTCTAGAGGCTGCTGGGCTGTTCAGTGTTCATCC	907
OY		838	-----ACAGAGTATTCCTTGTCTTCTTAAATCCATCCACCAACAACATAATTCACGTC	891
Db		908	ACATAAATACAGATATTCACACTTTATCTTACACAAACCCCAACGCCCACTCTCCACCTC	967
OY		892	ACTACTCCCAAGCCCTAC-----TTATCCCTGACTCCGCCCACT	936
Db		968	ACTACTCCCAAGCTGACCTTGAAGGCCCAAGATGATGACATCCCCCTGGCA	1027
OY		937	CACCCGACAGTGTATTGACTTGTGCAC-----	968
Db		1028	CAGACCCCAAGGCAATGTGTCTCACTGTACTCTGTGGCCAAAGATGGGTCCAGAAAGCC	1087
OY		969	-----CAGGCACTGATGATGGCTGACCTGGGGAGGACAGAGAAGCTGGGATAG	1023
Db		1088	CACCTTCAGGCACTAAGAGAGGCTGACACTGGGCGGAGAGCAAGAGACTGGGCTTAG	1147
OY		1024	GCCAGAAGTCCCACTGTAGGGGGAAGAGCTGGGGCAAGCTCTCCCTGGA----TC	1079
Db		1148	GCCAGAAGTCCCACTGTAGGGGGAAGAGCTGGGGCAAGCTCTCCCTGGA----TC	1206
OY		1080	CCTGTGATTTGAAAAATACTATTTTT	1108
Db		1207	CCTGTGATTTTAAACAGATATTATT	1235
<hr/>				
RESULT 8				
LOCUS	AC119115	138792 bp	DNA	linear HTG 18-JUL-2002
DEFINITION	Rattus norvegicus clone CH230-32ON23,*** SEQUENCING IN PROGRESS			
ACCESSION	AC119115			
VERSION	AC119115.2 GI:21746718			
KEYWORDS	HTG; PHASEL.			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 138792)

Murphy D.M., Adams C., Adio-Oduola B., Ali-Osman, F.R., Allen C.,  
Alsbrooks S.L., Amarantunge H.C., Are J.R., Ayele M., Banks T.,  
Bardaric J., Benton J., Blinag K., Blankenburg K., Bonni D.,  
Bouck J., Bowler S., Briteva M., Brown M., Bryant N.P.,  
Bunah C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,  
Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,  
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,  
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,  
Davila M.L., Davis C., Day-Carroll L., Dederich D.A.,  
DeLaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,  
Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,  
Eamhart C., Edgar D., Edwards C.C., Elhaj C., Escoto M.,  
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,  
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,  
Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,  
Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,  
Hernandez O., Hodgson A., Hognes M., Holloway C., Hollins B.,  
Homel F., Howard S., Huber J., Huiyk S., Hume J., Jackson L.E.,  
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,  
Karlssoon E., Kelly S., Khan U., King L., Korvah J., Kovar C.,  
Kratoch J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,  
Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louisseged H.,  
Lozado R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,  
Madeshwar M., Mapua P., Martin R., Martindale A., Martinez E.,  
Massey E., Mawhney E., McLeod M.P., Meador M., Mel G., Melker M.,  
Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,  
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,  
Nguyen N., Nickerson E., Nwokenkwo S., Ogum M., Okwuonu G.,  
Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,  
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,  
Rivers M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Savary G.,  
Scherer S., Scott G., Shen H., Shooshitari N., Sisson I.,  
Sodergren E., Sonake T., Sparks A., Stanley H., Stone H.,  
Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,  
Tansey J., Taylor C., Taylor T., Telitrod B., Thomas N., Thomas S.,  
Usmani K., Vasquez L., Vera V., Villalobos D., Vinson R., Wang O.,  
Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,  
Williams G., Williamson A., Wleczek R., Wooden S., Worley K.,  
Wu C., Wu Y., Wu Y.F., Zhou J., Zorilla S., Nelson D.,  
Weinstock G. and Gibbs R.

TITLE  
JOURNAL

Unpublished

2 (bases 1 to 138792)

Worley K.C.

Direct Submission

Submitted (25-Apr-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 138792)

Worley K.C.

Direct Submission

Submitted (18-Jul-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 14, 2002 this sequence version replaced gi.20303440.

## COMMENT

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GUXG

Center clone name: CH230-320N23

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 112720 bases at least Q40

Consensus quality: 116666 bases at least Q20

Consensus quality: 119165 bases at least Q20

\*\*\*\*\*  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1	1022:	contig of 1022 bp in length
1023	1122:	gap of unknown length
1123	2394:	contig of 1272 bp in length
2395	2494:	gap of unknown length
2495	3663:	contig of 1169 bp in length
3664	3763:	gap of unknown length
3764	4879:	contig of 1116 bp in length
4880	4979:	gap of unknown length
4980	6787:	contig of 1808 bp in length
6788	6887:	gap of unknown length
6888	8192:	contig of 1305 bp in length
8193	8292:	gap of unknown length
8293	9461:	contig of 1169 bp in length
9462	9561:	gap of unknown length
9562	11435:	contig of 1874 bp in length
11436	11535:	gap of unknown length
11536	13535:	contig of 2000 bp in length
13536	13635:	gap of unknown length
13636	17323:	contig of 3688 bp in length
17324	17423:	gap of unknown length
17424	20406:	contig of 2983 bp in length
20407	20506:	gap of unknown length
20507	23445:	contig of 2939 bp in length
23446	23545:	gap of unknown length
23546	25600:	contig of 2055 bp in length
25601	25700:	gap of unknown length
25701	28049:	contig of 2349 bp in length
28050	28149:	gap of unknown length
28150	30474:	contig of 2325 bp in length
30475	30574:	gap of unknown length
30575	35498:	contig of 4924 bp in length
35499	35598:	gap of unknown length
35599	39284:	contig of 3686 bp in length
39285	39384:	gap of unknown length
39385	41970:	contig of 2586 bp in length
41971	42070:	gap of unknown length
42071	45659:	contig of 3589 bp in length
45660	45759:	gap of unknown length
45760	50332:	contig of 4573 bp in length
50333	50432:	gap of unknown length
50433	54695:	contig of 4263 bp in length
54696	54795:	gap of unknown length
54796	58289:	contig of 3494 bp in length
58290	58389:	gap of unknown length
58390	63599:	contig of 5210 bp in length
63600	63699:	gap of unknown length
63700	70305:	contig of 6606 bp in length
70306	70405:	gap of unknown length
70406	76123:	contig of 5718 bp in length
76124	76223:	gap of unknown length
76224	84961:	contig of 8738 bp in length
84962	85061:	gap of unknown length
85062	93614:	contig of 8553 bp in length
93615	93714:	gap of unknown length
93715	103352:	contig of 9638 bp in length
103353	103452:	gap of unknown length
103453	110299:	contig of 6647 bp in length
110300	110399:	gap of unknown length
110400	116180:	contig of 5781 bp in length
116181	116280:	gap of unknown length
116281	125484:	contig of 9204 bp in length
125485	125584:	gap of unknown length

FEATURES	*	125585	138792:	contig of 13208 bp	in length
SOURCE			Location/Qualifiers		
		1.	138792		
		/organism="Rattus norvegicus"			
		/db_xref="taxon:10116"			
		/clone="CH30-320N23"			
BASE COUNT		34391	a 31996	c 31712	g 36193 t 4500
ORIGIN					

Query Match	42.9%	Score 501;	DB 2;	length 138792;
Best Local Similarity	83.6%	Pred. No. 7e-121;		
Matches 628; Conservative	0;	Mismatches 115;	Indels 8;	Gaps 5;

RESULT 9					
AX180714					
LOCUS	AX180714	898 bp	DNA	1 linear	PAT 06-AUG-2001
DEFINITION	Sequence 1 from Patent WO0145730.				
ACCESSION	AX180714				
VERSION	AX180714.1	GI:15132570			
KEYWORDS					

SOURCE ORGANISM	synthetic construct.
REFERENCE TITLE	synthetic artificial sequences. 1 (bases 1 to 898)
AUTHORS	Wiley, S.R.
JOURNAL	Tweak receptor
	Patent: WO 0145730-A 1 28-JUN-2001
	IMMUNEX CORPORATION (US)

CDS

```

/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MAYGSRPTSLALAFGLCLPWLQDGSATSSDRMKQIEDKTEILS
KYIENEIARIKKLIGERTSRSLGSRASLAEPAQEEIQAEDQDPSLNPQTEES
ODPAPFLNRLVPRPSAPKCGKRTPARAIAYHVHVRPGDDQAGADVGTSWGEA
RINSSPLRYNRNPGSEFIYTRAGLYLYICVHYDEGKAAYIKLIDLYDGLAIFCEE
FSNRAASLSIGQALCLCVSGSLALNRPSSSLIRITLPMAHLKAAPELTYFGELQVH"

```

Page 10

REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>
COMMENT	Contact: MGC help desk

## FEATURES

source	location/qualifiers
1. .1651	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20659 IMAGE:4766071"
/issue_type="Primary B-Cells from Tonsils"
/clone_lib="NIH-MGC-48"
/lab_host="DH10B-R"
/note="Vector: pOT87"
106. .510

```

BASE COUNT	ORIGIN
344 a	517 c 481 g 309 t

Query Match	35.1%;	Score 409.4;	DB 9;	Length 1651;
Best Local Similarity	73.7%;	Pred. No. 6.1e-97;		
Matches 685; Conservative	0;	Mismatches 171;	Indels 73;	Gaps 10;

298 AGCAGGTGTGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCC 357

|||||

	Db	713	ACCAGGCTGTGAGGGGACAGTGAAGTGGCTGGGAGGAGCAAGAAATCAACAGACTCCAGGCC	712
QY	358	TCGCGCTACGACGGCCAGATTTGGGGAATTTACAGTCAATCAGGGCTGGGCTCTACTACT	417	
Db	773	TCGCGCTACAAACCGCCAAATCGGGAGTTTATCTACACCGGGCTGGGCTCTACTACT	832	
QY	418	GTAAGTCAGAGTGCATTTTATGATGAGGAAAGAGCTGTCTACCTGAAGCTGGACTTGGAT	477	
Db	833	GTAAGTCAGAGTGCATTTTATGATGAGGAAAGAGCTGTCTACCTGAAGCTGGACTTGGAT	892	
QY	478	GAAAGGTGTGTGTGGGCTTGGCTGGCTGGAGAAATTTCTACGACACAGCAGACTCTCC	537	
Db	893	GGAATGTGTGTGGGCTTGGCTGGCTGGAGAAATTTCTACGACACAGCAGACTCTCC	952	
QY	538	TGGAGCCCAAGCTCCGTTTGTGCAAGTGTGGGTGTGGGTGTGGGGTGGGGCTTGC	597	
Db	953	CGGGCCCAAGCTCCGCTTGTGCAAGTGTGGGTGTGGGTGTGGGGTGGGGCTTGC	1012	
QY	598	CTTTGGATTCGACCCCTCCCTGGGCTCATTTAAGGCTGCCCTTCTCTAACTACTT	657	
Db	1013	CTGTGGATTCGACCCCTCCCTGGGCTCATTTAAGGCTGCCCTTCTCTAACTACTT	1072	
QY	658	TGAGACTTTTCAAGTTCAATGAGGGGCTTGTCTCTCCAGATTCTTAACTTTCCCTGG	717	
Db	1073	CGAGACTTTTCAAGTTCAATGAGGGGCTTGTCTCTCCAGATTCTTAACTTTCCCTGG	1132	
QY	718	CTTCAGAGACATCACACACCTCCCTACCCCAACCCCACTCTCCACCCCTC-GCTGC	776	
Db	1133	CTTCAGAGACATCACACACCTCCCTACCCCAACCCCACTCTCCACCCCTC-GCTGC	1187	
QY	777	CTTTGGTCCAGTCTGTCTCTCC--TCAAGGACGACAGAGCTTGTTCACATGTTTCCAT	834	
Db	1188	CTTTGGTCCAGTCTGTCTCTCC--TCAAGGACGACAGAGCTTGTTCACATGTTTCCAT	1247	
QY	835	TCC-----ACAGACGTATCTCTGTCTTTAAACATCCATCCACACACATTAACA	887	
Db	1248	TCCACATTAATTAACAGTATTCCTCACTTTATCTTAACAATCCCAACGCTCTCCA	1307	
QY	888	CCTCACTAGCTCCCAAGCCCTTAC-----TTATCCCTGACTCCCAAC	932	
Db	1308	CCTCACTAGCTCCCAAGCCCTTAC-----TTATCCCTGACTCCCAAC	1367	
QY	933	CAGTACCCGACAGAGTGTATTAATGACTTTGTGAC-----	968	
Db	1368	GCCACAGACCCCAAGAGGATTTGTCTCACTACTCTGTGGGCAAGAGTGGTCCAGAAG	1427	
QY	969	-----CAGGACTAGATGGGCTTGGACTGTGGGCAAGAGTGGTCCAGAAG	1019	
Db	1428	ACCCCACTTCAAGGCACTTAAGAGGGGGTGTGGACTGTGGGCAAGAGTGGTCCAGAAG	1487	
QY	1020	CTAGGCCAAGAGTTTCCCACTGTGAAGGGGGAAGAGCTGGGGACAAAGCTCTCCCTGA--	1077	
Db	1488	CTAGGCCAAGAGTTTCCCAATGTGAAGGGGGAAGAGCA--GACAAGCTCTCTCCCTGAGA	1546	
QY	1078	--TCCCTGGATTTTGTGA--AGATCACTATTTTATTTATTTGTGACAAAATGT--T	1130	
Db	1547	ATTCCCTGTGATTTTAAACAGATTTATTTTATTTATTTATTTGTGACAAAATGTGAT	1606	
QY	1131	AAATGATATTAAAGAGATTAATCATGA	1159	
Db	1607	AAATGATATTAAATAGATTAAGCATTA	1635	
RESULT 11				
AC127470				
Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15				
unordered pieces.				
AC127470				
AC127470.1 GI:2188686				
HTG; HTGS_PHASE1; HTGS_DRAFT.				
chimpanzee.				
SOURCE				
ORGANISM				
Pan troglodytes				









```

Db 37234 GTGCTGCCCCGCTGCGCTGCGTGGAGAAATTCACAGCCACGCGGCGAGTTCCCTCGGGCCC 37195
OY 545 CAGCTCCGTTTGTGCGAGGTGTCGGGCTGTCGGCGTGGCGGCGAGGCTCTTCCCTGGG 604
Db 37194 CAGCTCCGCTCTGCGAGGTGTCGGGCTGTCGGCGTGGCGGCGAGGCTCTTCCCTGGG 37135
OY 605 ATCCGACCCCTCCCTGGGCTGTCAGGTGTCGGCGTGGCGGCTCTTCCCTGGGCTG 664
Db 37134 ATCCGACCCCTCCCTGGGCTGTCAGGTGTCGGCGTGGCGGCTCTTCCCTGGGCTG 37075
OY 665 TTTCAGTTACATGAGGGGCTGCTGCTCCAGATTCCTTAACTTTCCCTGGCTCCAGG 724
Db 37074 TTCCAGTTACATGAGGGGCTGCTGCTCCAGATTCCTTAACTTTCCCTGGCTCCAGG 37018
OY 725 AGCATCACCACACGCTCCCTACCCGACCCGCTCCCTCCAGATTCCTTAACTTTCCCTGG 783
Db 37017 --CCTGACAGCTCTGTCGGGACCCGCTCCCTCCCTCCAGATTCCTTAACTTTCCCTGG 36960
OY 784 CCAGTCTGCTCTCTCC--TCAAAGCAGCCAGAGCTGTTTCATGTTTCCATTC-- 837
Db 36959 CCAGACCTGCGCCCTCCCTCTGAGAGCTGCTGCGGCTGTTTCATGTTTCCATTC-- 36900
OY 838 ---ACAGACGATCTCTGCTCTTAAACATCCATCCACCACTATCCACTCT 894
Db 36899 TAAATACAGTATTTCCCTCTTATCTTAACTCTCCACGCGCCACTCTCCACTCT 36840
OY 895 AGCTCCCAAAGCCCTAC-----TTATCCCTGACCTCCGACCCCACTCTAC 939
Db 36839 AGCTCCCAAAGCCCTAC-----TTATCCCTGACCTCCGACCCCACTCTAC 36780
OY 940 CCGACACAGCTGTTTATGACTTTTGAC-- 968
Db 36779 ACCCCAGAGGATGTTGTTACTGTCGTGCGGCAAGATGGTCCAGAGACCCAC 36720
OY 969 --CAGCAGTGAATGCTGAGCTGCTGTCGAGAGAGCAAGACTGAGGCTAGGCC 1026
Db 36719 TTCAGGACTAAGAGGGCTGAGCTGCGGAGAGCAAGAGACTGAGGCTAGGCC 36660
OY 1027 AGAATTTCCAACTGTGAGGGGAGAGAGCTGCGGCAAGCTCCCTGGA----TCCT 1082
Db 36659 AGGATTTCCAACTGTGAGGGGAGAG--AACAGACAAAGCTCTCCCTGAGAAATTCCT 36601
OY 1083 GTGATTTTGA--AGATTAATTTTATTTATTTATTTGACAAATGT--TAAATGGA 1137
Db 36600 GTGATTTTGA--AGATTAATTTTATTTATTTATTTGACAAATGTGATTAATGGA 36541
OY 1138 TATTAAGAGATTAATATGATTTCTCT 1166
Db 36540 TATTAAGAGATTAATATGATTTCTCT 36512

```

RESULT 13  
AC126921 153553 bp DNA linear HTG 10-JUL-2002  
LOCUS Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered  
DEFINITION pieces.  
AC126921  
VERSION AC126921.1 GI:21724098  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 153553)  
REFERENCE  
AUTHORS Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,  
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,  
Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
Marquies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D.,  
McCluskey, J.C., McDowell, J., Pasquillan, C., Pearson, R.,  
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantlirop, S., Thomas, J.W.,

Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 153553)  
REFERENCE  
AUTHORS Green, E.D.  
TITLE Direct Submission  
JOURNML Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
Groveom Circle, Gathersburg, MD 20877, USA  
COMMENT  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.zoehghri.nih.gov  
----- Project Information  
Center project name: ddi  
Center clone name: 045D24  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 14606 bases at least Q40  
Consensus quality: 147748 bases at least Q30  
Consensus quality: 148824 bases at least Q20  
Insert size: 151000; agarose-fp  
Insert size: 152453; sum-of-ctrls  
Quality coverage: 8.80x in Q20 bases; agarose-fp  
Quality coverage: 8.72x in Q20 bases; sum-of-ctrls  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces.  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
1 2313 2312: contig of 2312 bp in length  
\* 2413 2412: gap of unknown length  
\* 5842 5841: contig of 3429 bp in length  
\* 5842 5941: gap of unknown length  
\* 5942 8435: contig of 2494 bp in length  
\* 8436 8535: gap of unknown length  
\* 8536 15799: contig of 7264 bp in length  
\* 15800 15899: gap of unknown length  
\* 15900 25224: gap of 9325 bp in length  
\* 25225 25324: gap of unknown length  
\* 25325 32504: contig of 7180 bp in length  
\* 32505 32604: gap of unknown length  
\* 32605 40970: contig of 8366 bp in length  
\* 40971 41071: gap of unknown length  
\* 41071 36590: contig of 15320 bp in length  
\* 36591 56690: gap of unknown length  
\* 56691 73769: contig of 17079 bp in length  
\* 73770 73869: gap of unknown length  
\* 73870 90859: contig of 16990 bp in length  
\* 90860 90959: gap of unknown length  
\* 90960 111428: contig of 20469 bp in length  
\* 111429 111528: gap of unknown length  
\* 111529 153553: contig of 42025 bp in length.  
FEATURES  
source  
location/Qualifiers  
1..153553  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone="RP42-45D24"  
/clone\_11b="RP42"  
1..2312  
/note="assembly\_fragment"  
2413..5841  
/note="assembly\_fragment"  
clone\_end:77  
vector\_side:left"  
misc\_feature  
5942..8435  
misc\_feature







GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:27:27 ; Search time 1333.52 Seconds

(without alignments)

3449.164 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLEPRSLG.....PMALIKAPFLTYGILQVH 284

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n\_model -DEV=xlp  
-O=/cgn2.1/USPTO.spool/US09245198/runat\_06052003\_170239\_9628/app\_query.fasta\_1.846  
-DB=EST -QFMT=fastlap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09245198.ecgn\_1.1.2375 -runat\_06052003\_170239\_9628 -NCPV=6 -ICPV=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hmv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1087.5	73.3	731	13	B1871711
2	1060.5	73.4	728	13	B1870393
3	979	67.8	777	13	B1819200
4	962	66.6	834	13	B1766766
5	934	64.7	828	13	B1596681
6	913	63.2	918	12	BF577781
7	909.5	63.0	948	14	BQ707185
8	904	62.6	940	14	BQ884231
9	825	57.1	561	10	AW763237
10	809	56.0	963	14	BQ671259
11	765	53.0	1071	14	BQ692123
12	753	52.1	951	14	BQ674188
13	742.5	51.4	785	13	B1762908
14	738	51.1	587	13	BG686319
15	707	49.0	1033	11	AK020909
16	652	45.2	584	10	AW917574
17	648.5	44.9	698	13	B1906850
18	647	44.8	894	13	B1908274
19	645.5	44.7	474	13	B1965174
20	604	41.8	445	9	AA870722
21	598.5	41.4	471	9	AA221610
22	582.5	40.3	542	12	BF041509
23	582	40.3	650	12	BG404836
24	572.5	39.6	404	12	BF044430
25	562	38.9	360	10	BE654876
26	520.5	36.0	493	10	BE307031
27	506	33.0	558	12	BF821434
28	499	34.6	538	13	BA484863
29	493	34.1	298	14	BM688946
30	476	33.0	441	13	B1967060
31	466	33.3	315	12	BF466521
32	441	30.5	465	9	A1091441
33	440	30.5	581	13	B1738634
34	421	29.2	345	14	R55379
35	420	29.1	278	10	AA417023
36	403	27.9	1187	14	BQ053284
37	395.5	27.4	571	12	BF073881
38	391	27.1	639	12	BB642326
39	388.5	26.9	1926	14	BM906056
40	363	25.1	400	9	A1152313
41	362	25.1	342	9	AA637970
42	353	24.4	910	12	BG110063
43	352	24.4	456	13	B1966255
44	347	24.0	487	10	AW320117
45	320	22.2	533	10	BE628951

#### ALIGNMENTS

RESULT 1  
LOCUS B1871711 731 bp mRNA linear EST 11-OCT-2001  
DEFINITION 603395825F1 NIH\_MGC\_90 Homo sapiens CDNA clone IMAGE:5405478 5',  
ACCESSION B1871711  
VERSION B1871711.1 GI:16045386  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 731)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



```

OY 104 SerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 123
    |||||||
Db 122 TCGGACAGTGAATCCCGACAGACAGAAAGACCAGATCCGGCCCTTCTCGAACCAGCTA 181
OY 124 ValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla 143
    |||||||
Db 182 GTTGGCCCTCGGAGAGGACCTTAAGCCGGAAACACGGGGCTCGAAGACGATCGCA 241
OY 144 AlaHisThrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
    |||||||
Db 242 GCCCTTTATGAGATTCATCCACGACCTGACACGAGACGGACGACGAGCTGTGGCCG 301
OY 164 ThrValSerGlyTrpGluGluAlaArgIleAsnSerSerProLeuArgTyrAsnArg 183
    |||||||
Db 302 ACAGTGAATGCTGGAGAGAGACCCAGATCAACAGCTCCAGCCCTCTGCTACACACCGC 361
OY 184 GlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrIleuTyrCysGlnValHis 203
    |||||||
Db 362 CAGATCGGGAGTATTATAGTCAACCGGGCTGGCTCTACTACTGACTGACGAGTCAC 421
OY 204 PheAspGluGlyLysAlaValTyrIleuLysLeuAspLeuValAspGlyValLeuAla 223
    |||||||
Db 422 TTTGATGAGGGGAGAGCTGTCTACTGAGCTGAGCTGCTGCTGATGCTGTGGCC 481
OY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerIleuGlyProGlnIleuArg 243
    |||||||
Db 482 CTGGCCCTGCTGGAGGAATTCACACGACCTGGCCAGTTCCTCGGGCCCAAGCTCCGC 541
OY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263
    |||||||
Db 542 CTCTCCAGAGTGTCTGGGCTGTGGCCCTCGCGCCAGAGTCC-TCCCTGGAGATCCGCAC 600
OY 264 LeuProTrpAlaHisIleuLysAlaAlaArgProPheLeuThrTyrPheGlyLeuArgIleVal 283
    |||||||
Db 601 CTCCTCCGGGCCCATCTCAAGGCTGC-CCCTTCTCTACTACTGTGGACTCTTCAGAGTT 659
OY 284 His 284
    |||
Db 660 CMC 662

RESULT 3
LOCUS B1819200 777 bp mRNA linear EST 04-OCT-2001
DEFINITION 603034614P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',
    mRNA sequence.
ACCESSION B1819200
VERSION B1819200.1 GI:15930750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
    Email: cgabbs-remail.nih.gov
    Tissue Procurement: Life Technologies, Inc.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNLN at:
    http://image.lnl.gov
    Plate: L14M11437 row: 1 column: 03
    High quality sequence stop: 759.
FEATURES
    Source
        1..777
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5175698"

```

```

/clone.lib="NIH_MGC_115"
/lab.host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-sports; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC library."

BASE COUNT 140 a 268 c 215 g 154 t
ORIGIN

Alignment Scores:
Pred. No.: 5,336-83 Length: 777
Score: 979.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.80% Indels: 0
DB: 13 Gaps: 0

US-09-245-198A-4 (1-284) x B1819200 (1-777)

OY 97 AlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro 116
    |||||||
Db 14 GCAGAGGAGACAGACAGACCCGTCGGAACTGAATCCCAACAGAAAGCCAGATTCCT 73
OY 117 AlaProPheLeuAsnArgLeuValArgProArgSerAlaProLysGlyArgLysThr 136
    |||||||
Db 74 GGGCTTTCCTGAACCGACTAGTGGCCCTCCAGAAAGTGCACCTTAAGCCGGAAAAACA 133
OY 137 ArgAlaArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGly 156
    |||||||
Db 134 CGGGCTCGAAGAGCGATCCACGCCCATGTATGAAGTTCATCCACAGCTGAGACGCGA 193
OY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSer 176
    |||||||
Db 194 GCGAGGACAGGTGTGTGACGCGAGAGTGTGTGGAGAGAAAGCCAGAAATCAACAGCTCC 253
OY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196
    |||||||
Db 254 ACCCTCTGCGCTACACCGCCAGATCGGGGAGTTATATGTCACCGGGCTGGCTTAC 313
OY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216
    |||||||
Db 314 TACCTGTAAGTGCAGAGTGCATTTGATGAGGGGAAGGCTGTCTACTGAACTGGACTTG 373
OY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAspSer 236
    |||||||
Db 374 CTGGTGGATGAGTGTCTGCGCCCTGCGCTCCCTGAGGAATTCACGCCCTCGCGAGCT 433
OY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256
    |||||||
Db 434 TCCCTCGGGGCCACCTCGCCCTGCGCAGAGTGTCTGGGCTTTGGCCCTGGCGCAGGG 493
OY 257 SerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThr 276
    |||||||
Db 494 TCCCTCGGGGATCGGACCTCCCTCGGGGCCATCTCAAGGCTGCCCCCTCTTACAC 553
OY 277 TyrPheGlyLeuPheGlnValHis 284
    |||||||
Db 554 TACTTCGGACTTCTCCAGGTTTAC 577

RESULT 4
LOCUS B1766766 834 bp mRNA linear EST 25-SEP-2001
DEFINITION 603056866P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
    mRNA sequence.
ACCESSION B1766766
VERSION B1766766.1 GI:15758344
KEYWORDS EST.

```



SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
AUTHORS NIH-MGC <http://mgi.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Plate: LLM11517 row: c column: 18  
High quality sequence stop: 772.  
Location/Qualifiers  
1. 834  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5206217"  
/clone\_1lb="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC library."  
BASE COUNT 154 a 287 c 223 g 170 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,46e-81 Length: 834  
Score: 962.00 Matches: 192  
Percent Similarity: 98.97% Conservative: 0  
Best Local Similarity: 98.97% Mismatches: 1  
Query Match: 66.62% Indels: 2  
DB: 13 Gaps: 0

US-09-245-198a-4 (1-284) x B1766766 (1-834)  
OY 92 G1ng1ug1ueuVala1ag1ug1uaspG1naSPROSeR1ueuAsnPRoG1nThG1u 111  
DB 3 CAGAGAGAGCTGGTGGCAGAGAGACAGACCCGCGAGACATGATCCCAAGAGAA 62  
OY 112 G1uSerG1naSPROa1aPhe1euaN1aR1eua1aRPRoARgARSeR1aPro 131  
DB 63 GAAAGCCAGAGATCTGGCCCTTCCTGTAACCGACTAGTGGCTCCAGAGAGGACCT 122  
OY 132 LysG1ArG1ySThArG1a1aArG1a1a1eAla1a1a1sTyrG1uVal1H1sPRoArG 151  
DB 123 AAAGGCCGGAACAGCGGCTCGAAGAGCATGCAACCATATATGAAGTTCATCCACA 182  
OY 152 Pro-G1yG1naSPe1y1a1aG1y1a1aSPeG1yTh1a1aSPeG1yTTPeG1uVal 171  
DB 183 CCGTGGACAG 242

OY 171 aArG11eAnS1eRSeRSeRPro1euaRg1yR1aRnArG1n11eG1y1u1u1h1e1e1a1h 191  
DB 243 CAGAAATCAACAGCTCCAGCCCTTCGCGCTACAAACCGCAATCGGAGATTATAGTCAC 302  
OY 191 rArG1aG1y1e1u1yT1r1e1u1yT1r1e1u1yT1r1e1u1yT1r1e1u1yT1r1e1u1y 211  
DB 303 CCGGGGCTGGGCTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 362

OY 211 r1e1u1y1e1u1aSP1e1u1e1u1aSPG1y1a1e1u1a1a1e1u1aArG1ySe1u1u1h1e1e 231  
DB 363 CCGAAGCTGAGACTTGCTGGTGGATGCTGCGCCCTCGCGCTCGAGAGAAATTC 422  
OY 231 rAlaThAla1a1e 251  
DB 423 AGCCACAGCGGCGAGTCCCTCGGGCCCGAGCTCCGCTCTGCAAGAGTCTGGCGT 482  
OY 251 uAla1e1u1aRPRoG1ySeRSeRSeRSeRSeRSeRSeRSeRSeRSeRSeRSeRSeRSeR 271  
DB 483 GGGCCCTGGCCAGAGTCC-TCCTGGCGAGATCCGACCCCTCCCTGGGCCCATCAAGC 541  
OY 271 aAlaPRoPhe1e1u1Th1r1e1r1e1e1y1e1u1Phe1e1u1a1H1s 284  
DB 542 TGGCCCTTCCTCACACTTCCGAGCTCTCCAGGTTCCAC 581

RESULT 5  
B1596681 828 bp mRNA linear EST 07-SEP-2001  
LOCUS 603243254F1 NIH\_MGC\_96 Homo sapiens CDNA clone IMAGE:5285892 5',  
DEFINITION mRNA sequence.  
ACCESSION B1596681  
VERSION B1596681.1 GI:15489620  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 828)  
AUTHORS NIH-MGC <http://mgi.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Miklos Palovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Plate: LLM11722 row: k column: 13  
High quality sequence stop: 776.  
Location/Qualifiers  
1. 828  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5285892"  
/clone\_1lb="NIH\_MGC\_96"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC library."  
BASE COUNT 155 a 278 c 223 g 172 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,13e-78 Length: 828  
Score: 934.00 Matches: 179  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.44% Mismatches: 0  
Query Match: 64.68% Indels: 0  
DB: 13 Gaps: 0

US-09-245-198a-4 (1-284) x B1596681 (1-828)  
OY 155 a 278 c 223 g 172 t





REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM2466 row: n column: 17  
 High quality sequence start: 24  
 High quality sequence stop: 550.  
 Location/Qualifiers

FEATURES  
 source  
 1..948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH-MGC\_113"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH-MGC Library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,95e-76 Length: 948  
 Score: 909.50 Matches: 179  
 Percent Similarity: 98.358 Conservative: 0  
 Best Local Similarity: 98.358 Mismatches: 2  
 Query Match: 62.984 Indels: 2  
 DB: 14 Gaps: 1

US-09-245-198a-4 (1-284) x BQ707185 (1-948)

QY 103 ProSerGluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeuAsnArg 122  
 Db 3 CCGTCGGAATCGAATCCCGACAGACAGAGAACGACAGATCTCGCCTTCGTAACCGA 62  
 QY 123 LeuValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIle 142  
 Db 63 CTGATTCGGCTCGAAGAACTGACCTAAAGCGCGGAAACACGCGCTCGAACAGCGATC 122  
 QY 143 AlaAlaHisArgLysValHisProArgProGlyGlnAspGlyAlaGlnAlaValAsp 152  
 Db 123 GCGAGCCATTATGAGTTCACGACGACCTGGACAGACGAGAC---GCAGGTGTGAC 178  
 QY 163 GlyThrValSerGlyTyrGlnGluAlaArgIleAsnSerSerSerProLeuArgTyrAsn 182  
 Db 179 GGGACAGTGAAGTGGTGGGAGGAGGAGACGACATCAACAGCTCCAGCCCTCGCCTACAC 238  
 QY 183 ArgGlnIleGlyLysPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 202  
 Db 239 CGCCACATCGGAGATTATAGTCACCCGCGGTGGCTCTACTACTGCTACAGTGC 298  
 QY 203 HisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeu 222  
 Db 299 CACTTGTATGAGGAGGAGGCTGTCTACCTGAACCTGACTGTGTGGTGTGTGCTG 358  
 QY 223 AlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeu 242  
 Db 359 GCCCTCGCGCTCGGAGGAGATTCACGACCTGCGGCGGCTTCCTCGGCGCCACAGTC 418  
 QY 243 ArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArg 262

Db 419 CCGCTCGCAGAGTGTCTGGGCTGTGGCCCTCGGCGCAGAGCCCTCGGAGATCCGC 478  
 QY 263 ThrLeuProTAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln 282  
 Db 479 ACCCTCCCTCGGCGCATCTCAAGGCTGCCCTCTCTACCTGACCTTCACAG 538

QY 283 ValHis 284  
 Db 539 GTTCAC 544

RESULT 8  
 BQ884231 940 bp mRNA linear EST 16-AUG-2002  
 LOCUS BQ884231  
 DEFINITION AGENCOURT 8682031 lupskl.scfatic.nerve Homo sapiens cDNA clone  
 IMAGE:6197488 5', mRNA sequence.  
 ACCESSION BQ884231  
 VERSION BQ884231.1 GI:22276239  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 940)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM13607 row: 1 column: 17  
 High quality sequence stop: 453.  
 Location/Qualifiers

FEATURES  
 source

1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Lupskl.scfatic.nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TGCACCCACGCGTCCG-3' and  
 5'-GACTGCTTGTAGTACGCGGAGCGGCGCCCT(15)-3'. Site selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 9.73e-76 Length: 940  
 Score: 904.00 Matches: 183  
 Percent Similarity: 97.878 Conservative: 1  
 Best Local Similarity: 97.348 Mismatches: 0  
 Query Match: 62.608 Indels: 4  
 DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x BQ884231 (1-940)

QY 101 GlnAspProSerGluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeu 120  
 |||||||

Db 1 CAGGACCCGTCGAGTGAATCCAGACAGAGAAAGCCAGATCTCGCTTCTCTG 60  
QY 121 AsnArgLeuValArgProArgSerAlaProIysGlyArgLysThrArgAlaArg 140  
Db 61 AACGACAGTTCGGCTCCGAGAGTCACCTAAAGCCGGAACACGGGGTCGAGA 120  
QY 141 AlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
Db 121 GCGATCCGACCCCTTATGAAATTCATCCAGCAGCTGACAGCAGGAGCCAGCAGT 180  
QY 161 ValAspGlyThrValSerGlyTyrPgluLualaArgIleAsnSerSerProLeuArg 180  
Db 181 GTGACGGGACAGTGAAGGCTGGAGAGACCAAGATCAACAGCTCCAGCCCTCTGCGC 240  
QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 241 TACAAACCCGACATCGGGAGATTATATCTACCCGCGGCTGAGCTTACTACTGACGT 300  
QY 201 GlnValHisPheAspGlyLysAlaValAlaTyrLeuLysLeuAspLeuValAspGly 220  
Db 301 CAGGTGCACCTTGATGAGGAGGAGAGGCTGTCTACTGAAAGCTGAGACTTCTGTGATGCT 360  
QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyPro 240  
Db 361 GTGCTGGCCCTGCTGCTGCTGAGAGATTTCTACGACCTGCGGCGAGTTCTCTGCGGCC 420  
QY 241 GlnLeuArgLeuCys-GlnValSerGlyLeuLeuAlaLeuArgProGlySer-SerLeuA 260  
Db 421 CAGTCCGCTCTCTCCAGGAGGTCTGGGCTGTGGCCCTGCGGCGAGGGCCCTCTCCCTGC 480  
QY 260 rGileArgThrLeuProTrpAlaHisLeuLysAlaAla-PropheLeuThrTyr-Phel 279  
Db 481 GGAATCCGACCCCTCCCGGGGCCCATCTCAAGGCTGCGCCCTTCTACTACTTTCGG 540  
QY 279 yLeuPheGlnValHis 284  
Db 541 ACTCTCCAGGTCAC 556

RESULT 9  
AW763237 561 bp mRNA linear EST 04-MAY-2000  
LOCUS  
DEFINITION  
un70409.y1 NCI-CGAP\_Mam3 Mus musculus cDNA clone IMAGE:315563 5'  
similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS  
; mRNA sequence.  
ACCESSION  
AW763237  
VERSION  
AW763237.1 GI:7695174  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 561)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Other-ESTs: un70409.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
image.liml.gov/image/html/lresources.shtml

MG1:1058389  
Seq primer: -40RP from Gldco  
High quality sequence stop: 433.  
Location/Qualifiers  
1..561  
/organism="Mus musculus"

/strain="129, C57BL/6J, FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:315563"  
/clone\_lib="NCI-CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI. Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

BASE COUNT 108 a 156 c 194 g 100 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.52e-68 Length: 561  
Score: 825.00 Matches: 161  
Percent Similarity: 90.91% Conservative: 9  
Best Local Similarity: 86.10% Mismatches: 17  
Query Match: 57.13% Indels: 0  
DB: Gaps: 0

US-09-245-198A-4 (1-284) x AW763237 (1-561)

QY 41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuValProLeuAla 60  
Db 1 AGCCGAGGCGGAGGGGGGCGGGGGGAGCGGGGACCGCCCTGCTGCTGGTG 60  
QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValAlaSerLeuGly 80  
Db 61 CTGACCTCGGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 120  
QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAlaGluLasp 100  
Db 121 AGCTGGGACAGCTGTCTGCCAGAGGCTTCTCAGAGAGAGCTACAGAGAGAGCCG 180  
QY 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProheLeu 120  
Db 181 CGGGAGCCCTCGAAGCTAATCCAGACAGAGGAGAAACCGAGATGTGACTTCTTG 240  
QY 121 AsnArgLeuValArgProArgSerAlaProIysGlyArgLysThrArgAlaArg 140  
Db 241 GACACACTAGTCCGGCTCGAAGAGTCTCTTAAGCCGGAAGCCGCTCGCCGA 300  
QY 141 AlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
Db 301 GCTATTGCAGCCCATTTAGAGGTTTCTCTGCGGACAGAGAGAGGAGCAGCAGT 360  
QY 161 ValAspGlyThrValSerGlyTyrPgluLualaArgIleAsnSerSerProLeuArg 180  
Db 361 GTGATGGGACAGTGAAGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTGCGC 420  
QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 421 TACGACCCGACATTTGGGGAATNTACAGTCAAGGCTGGCTTACTACTGACGT 480  
QY 201 GlnValHisPheAspGlyLysAlaValAlaTyrLeuLysLeuAspLeuValAspGly 220  
Db 481 CAGGTGCACCTTGATGAGGAGGAGAGGCTGTCTACTGAAAGCTGCTGTGAAAGCT 540  
QY 221 ValLeuAlaLeuArgCysLeu 227  
Db 541 GTGCTGGCCCTGCGCTGCTG 561

RESULT 10  
BO671259 963 bp mRNA linear EST 15-JUL-2002  
LOCUS  
DEFINITION  
ACENCDUTL\_8303564 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6247716  
5', mRNA sequence.  
ACCESSION  
BO671259  
VERSION  
BO671259.1 GI:21782093



QY	156	GINALAGLYVALASGLYFhValSerGIYTPGIUGIUALATGILLeasSerSerSer	177
Db	72	CAGGCAAGGTGTGACAGGAGACGTGAGTGGCTGGGAGGAACGAGATACAGAGCTCCAGC	131
QY	178	ProLeuArGTYrAsnArGlnIleGIyLuphelleValIThArGAlaGIyLeuTYr	197
Db	132	CCTTCGGCCTACAAACCCGACAGTCGGGGAGTTTAAAGTCACCCGGGCTGGGCTCTACTAC	191
QY	198	LeuTYrCySGlnValHisPheAspGIUGIyLysAlaValIYrLeuLysLeuAspLeu	217
Db	192	CGTACTCTCAGAGTCACATTGTATGAGGGGAAGCGCTGCTACCTGAACCTGAGACTGCTG	251
QY	218	ValAspGIyValLeuAlaLeuAArgySLeuGIUGIyLupheserAlaThAlAlaSerSer	237
Db	252	GGGATATGGTGTCTGGCCCTGGCTGCTGGAGGAATTCCTAGCCACATCGGCGAGTTCC	311
QY	238	LeuGIyProGlnLeuAArgLeuCySGlnAlaSerGIyLeuAlaAlaLeuAArgProGIySer	257
Db	312	CTCGGGGCCCAAGCTCGCGCTCTGCCAGAGTGTCTGGGCTGTGGCCCTCGGGCCAGAGTCC	371
QY	258	SerLeuArGIleArYFhLeuProTPAlaHisLeuLysAlaAlaProPheLeuThrTYr	277
Db	372	TCCCGCCGAGATCGCACCTCCCTCGGGCCCATCTCAAGCGTGGCCCTCTCTCACTAC	431
QY	278	PheGIyLeuPheGlnValHis	284
Db	432	TTCCGAGCTCTTCCAGGTTTAC	452
RESULT 12			
LOCUS	B0674188	951 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT.8354100 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275664		
ACCESSION	B0674188		
VERSION	B0674188.1	GI:21785022	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: c9apbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMU at:		
	http://image.lnl.gov		
	Plate: LICM2459 row: d column: 01		
	High quality sequence stop: 335.		
FEATURES	Location/Qualifiers		
source	1..951		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6275664"		
	/clone_1lb="NIH_MGC_102"		
	/tissue_type="epidermoid carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: salivary gland; Vector: pOT87; Site_1: XhoI;		
	Site_2: EcoRI; cDNA made by oligo-dT priming.		
	Directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGGACGAG(G).		
	by Ling Hong in the laboratory of Gerald M. Rubin		
	(University of California, Berkeley) using ZAP-cDNA		
	synthesis kit (Stratagene) and Superscript II RT (Life		
	Technologies). Note: this is a NIH_MGC library."		
BASE COUNT	165 a 286 c 315 g 184 t 1 others		
ORIGIN			

Alignment Scores:	2.35e-61	Length:	951
Pred. No.:	753.00	Matches:	152
Score:	95.62%	Conservative:	1
Percent Similarity:	95.00%	Mismatches:	5
Best Local Similarity:	52.15%	Indels:	2
Query Match:	14	Gaps:	0

  

US-09-245-198A-4 (1-284) x B0674188 (1-951)
QY 124 VALARPROARGARGSERALAPROLYSGIARGLYTHRARGALARGARALALEA 143
DB 2 GTTGGCGCTCGAGAAAGTGACACTAAAGCCGGAAACACGGGGCTCGAAGACGATCGCA 61
QY 144 ALAHISTYRGUVALHISPROARPROGLYGINASPGYALAGINALAGLYVALSPGY 163
DB 62 GCCCATTTATGAAAGTTCATCCACGACCTGGACAGGAGCGCCAGACGATGTGGACGGG 121
QY 164 THRVALSERGATYTRGUGLUNALARGILEASNSRSESERPROLEUARGTYRANRG 183
DB 122 ACAGTGAATGGCTGGAGAGAACCCAGATCAACAGCTCCAGCCCTCTGCGCTACACACCC 181
QY 184 GINILEGUGLUPHEILEVALTHARAGIAGLYLEUTYTYRILEUTYCYSGINALHIS 203
DB 182 CAGATCGGGGAGATTATATAGTACCCGGGGCTGGCTCTACTACTGTAAGTGAGGCAC 241
QY 204 PHEASPGUGLYLYSVALAVALTYRLEUTYLSLEASAPLEULEVALASPGLYVALLEUALA 223
DB 242 TTTGATGAGGGGAAAGGCTGTCTACTCGAAGCGGACCTGCTGTGTGATGTGTGGCTGGCC 301
QY 224 LEUARGCYLSLEUGLIGLPHESERLATHRRALALASERSELEUGLYPROGINTLEUARG 243
DB 302 CTGGCTGCTCGTGGAGGAATCTCACCCACTGGGGCCAGTCCCTGGGGGCCACGCTCCGC 361
QY 244 LEUCYSGINVALSERGILEULEUALALEUARG-PROGLYSERSELEUARGILEARGTH 263
DB 362 CTCTCCAGAGTGTCGGGCTGTGGCCCTCGCGGCGACAGGCTCTCCCTGGGATGCCAC 421
QY 263 RLEUPROTTHRALHISLEULYSVALALA-PROPHLEUTHRYRPHGELYLEUPHE 281
DB 422 CTTCCCGGGGGGCCATTTAAAGGTGGCCCTCTTCTCTCACATTTTGGAAATTTT 477

  

RESULT 13	785 bp	mRNA	linear	EST 25-SEP-2001
LOCUS B1762908				
DEFINITION 60304796661 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188168 5',				
ACCSSION B1762908		mRNA sequence.		
VERSION B1762908.1	GI:15754486			
KEYWORDS EST.				
SOURCE human.				
ORGANISM Homo sapiens				
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE 1 (bases 1 to 785)				
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.				
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished (1999)				
Contact: Robert Strausberg, Ph.D.				
Email: cga@bbs-remail.nih.gov				
Tissue Procurement: Life Technologies, Inc.				
cDNA Library Preparation: Life Technologies, Inc.				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
DNA Sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be				
found through the I.M.A.G.E. Consortium/LNLN at:				
http://image.lnl.gov				
Plate: L1AM11470 row: c column: 17				
High quality sequence stop: 647.				
Location/Qualifiers				
1. .785				
source				
organism="Homo sapiens"				

DB:	Sequence	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
US-09-245-198A-4 (1-284) x BT62908 (1-785)		1,78e-60	785	195	5	36	22
Alignment Scores:							
Pred. No.:		742.50					
Percent Similarity:		77.52%					
Best Local Similarity:		75.58%					
Query Match:		51.42%					
DB:		13				7	
BASE COUNT	124 a 228 c 274 g 159 t						
ORIGIN							
	/db_xref="taxon:9606" /clone="IMAGE:5188168" /clone_id="NIH_MGC_116" /lab_host="DH10B" /note="Organ: pooled colon, kidney, stomach; Vector: PCWV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); source anonymous pool of 3 colons, age 26 yo male, 49 female, 71 yo male colon; 46 yo male kidney, and pool stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by Gruber (Invitrogen). Research Genetics Tracking code 023. Note: this is a NIH_MGC Library."						
25	glaValaValaArgGlnAlaGlnProProAlaProMetAla-AlaArgArgSerGlnArgAr	44					
29	ggCGGTGATGTCACATGCACACACCCCGCCCGCCATTTCCTGCGCGTGGAGCCAGAGCG	88					
44	gargGlyArgArg-Gly-GluProGly---ThrAlaLeuLeuValProLeuAlaLeuGly	62					
89	gAGGGGGCGCGGTGGTGGAGACCTGGGCGACCTGGCTGTGGGTCCTGGCTGCTGCTGG	148					
63	---LeuGlyLeuAla-LeuAlaCysLeu-----GlyLeuLeuLeuAlaVala	77					
149	TGCGTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	208					
77	lSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGln-GluGluLeuVala	97					
209	GAGCTGGGC-----ATCGCTGTCGCCCGCCAGAGACCTGCCAGAGAGAGACTGTTG	262					
97	laGlu---GluAspGlnAspProSer---GluLeuAsnProGlnThrGluGluSerGlna	115					
263	GCAGAGTGAATGACACAGAGACCTGTCTGGAACTGATCCCGACAGAGAAGAACGACAG	322					
115	sppProAlaProPhleLeuAsnArg-LeuValaArgProArgArgSerAlaProPolyGlyArg	134					
323	ATTCCTGGGCTTCTCTTAACCTGACTAGTTGGGCTTGGCGAGAAAGTGACCTAAAGGCCG	382					
135	LySThrArgAlaArgArgAlaLeuAlaAlaHisTyrGluValaHisProArgProGlyGln	154					
383	AAAACACGGCGCTGAAAGAGCATGGCAGCCATTATATAATTCATCCAGACACTGGACAG	442					
155	AspGlyAlaGlnAlaGlyValaAspGlyThrVal-SerGlyTyrPoluGlnAlaArgIleAs	174					
443	GACGAGGAGCGACAGGTGTGGAGGAGACATTGAGTGGCTGGGAGAGAACCGAATTCMA	502					
174	nSerSerSerProLeuArgTyrAsnArgGlnIleGlyGluPhleIleValaThrArgAlaG	194					
503	CAGCTCCAGCCCTCTGGCGTCAACAACCGCCAGATCGGGGAGCTTATATATGTCACCCG	562					
194	ylLeuTyrTyrLeuTyrCysGlnValaHisPheAspGlnGlyAlaValaTyrLeuGlySte	214					
563	GCTTACTAGTACTGTACTGTCTAGGTGCACTTTGATGAGGGGAGGCTGTCTACTGAAGCT	672					
214	uAspLeuLeuValaAspGlyAlaLeuAla---LeuArgCysLeuGlnGluGluPheSerAlaTh	233					
623	GGACTTGCTGGTGGATGATGTGCTGGGCGCGTGGCTTGGAGAGACATTTCTCAGAGCAC	682					
223	rAlaAla-SerSerLeu-GlyProGlnLeuArgLeuCysGlnValSer-GlyLeuLeu-A	252					

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	663	TCGGCGGAGTTCCTCGGGGCCCCAGTCGACCTCTCTCCACAGTCTCTGGGCTTTGGG	742					
Db	252	1aleuAurProGlySerSerLeuArgIleArg	262					
Db	743	CCTTGGGGGCGAGGCTCTCGCTCGCATGCCG	774					
RESULT 14								
	BG686319	587 bp	mRNA	linear	EST	01-MAY-2001		
	602638232p1 NIH_MGC_48	Homo sapiens	CDNA	clone	IMAGE:4766071.5'			
DEFINITION								
LOCUS	BG686319							
ACCESSION	BG686319							
VERSION	BG686319.1	GI:13917716						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
TITLE	NIH-MGC http://mhc.nci.nih.gov/.							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cgapbs-remail.nih.gov							
	Tissue Procurement: Louis M. Staudt, M.D., Ph.D.							
	CDNA Library Preparation: Ling Hong/Rubin Laboratory							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNLN at:							
	http://image.lnl.gov							
	Plate: L16M1625 row: p column: 08							
	High quality sequence stop: 587.							
FEATURES								
SOURCE	Location/Qualifiers							
	1..587							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:4766071"							
	/clone_1ID="NIH_MGC_48"							
	/tissue_type="primary B-cells from tonsils (cell line)"							
	/lab_host="DH10B (phage-resistant)"							
	/note="Organ: B-cells; Vector: pOTB7, site_1: XhoI;							
	site_2: EcoRI; CDNA made by oligo-dT priming.							
	Directionally cloned into EcoRI/XhoI sites using the							
	following 5' adaptor: GGCACGAG(G). Size-selected >500bp							
	for average insert size 1.8kb. Library constructed by Ling							
	Hong in the laboratory of Gerald M. Rubin (University of							
	California, Berkeley) using ZAP-cDNA synthesis kit							
	(Stratagene) and Superscript II RT (Life Technologies).							
	Note: this is a NIH_MGC Library."							
BASE COUNT	108 a	187 c	205 g	87 t				
ORIGIN								
Alignment Scores:								
Pred. NO.:	3.12e-60	Length:	587					
Score:	738.00	Matches:	152					
Percent Similarity:	97.44%	Conservative:	0					
Best Local Similarity:	97.44%	Mismatches:	1					
Query Match:	51.11%	Indels:	3					
DB:	12	Gaps:	1					
US-09-245-198A-4 (1-284) x BG686319 (1-587)								
QY	9	Ser1AaGagGluProLeuProArgSerLeuGlySerArgAspGlyAlaValArg	28					
Db	17	TCGGCGGCGGCTCCCTCCCGAGTCCTCGGGTCCCGAGGGGGGGCGGTAGG	76					
QY	29	GlnAlaGlnProPheAlaProMetAlaIleArgArgSerGlnArgArgGlyArgArg	48					
Db	77	CAGCAGACAGCCCCCGCCCATGGCCCGCTGGAGACAGAGCGGAGGGGGCGCCG	136					
QY	49	GlyGlnProGlyThrAlaLeuLeuValProLeuAlaIleGlyLeuGlyLeuAla-LeuAl	68					





gene 1. .1033  
/gene="tnsf12"  
<1. .433  
/gene="tnsf12"  
/note="data source:MCD, source key:MGI:1196259,  
evidence:ISS  
putative  
tumor necrosis factor (ligand) superfamily, member 12"

/codon\_start=2  
/protein\_id="BAB32249.1"  
/db\_xref="GI:12861641"  
/db\_xref="MGI:1196259"  
/translation="1AAHEVHPRGQDAQAGVDGVSQWEETKINSSPLRYDROI  
GEFTVRAGLVYLCOVHDEGAAVYKLDLVNGVIALRCLEEFSAATASSPGPOLR  
LCOVSGILPLRPSSSLRIRTLPPAHLKAAPFLTYFGLFOVH"  
BASE COUNT 228 a 300 c 248 g 256 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.:	6.36e-57	Length:	1033
Score:	707.00	Matches:	135
Percent Similarity:	96.50%	Conservative:	3
Best Local Similarity:	94.41%	Mismatches:	5
Query Match:	48.96%	Indels:	0
DB:	11	Gaps:	0

US-09-245-198a-4 (1-284) x AK020909 (1-1033)

QY 142 ILAAlaAlaHisTyrGluValHisProArgProGlyValAspGlyAlaGlnAlaGlyVal 161  
|||  
Db 2 ATTCAGACCCCATTAATGAGTTCATCCCGGCAGCAGATGAGCACAAGCAGGTGTG 61  
|||  
QY 162 AspGlyThrValSerGlyTyrGluAlaArgIleAsnSerSerSerProLeuArgTyr 181  
|||  
Db 62 GATGGGACAGTGAAGTGGCTGGAGAACCAAAATCAACAGCTCCAGCCCTGCGCTAC 121  
|||  
QY 182 AsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGln 201  
|||  
Db 122 GACCGCCAGATTGGGCAATTACATCATCAGGCGCTGCTACTACTGCTACTGTCAG 181  
|||  
QY 202 ValHisPheAspGluGlyLysAlaValTyrLeuLysAspLeuValAspGlyVal 221  
|||  
Db 182 GTCCACTTATGAGGAAAGGCTGTCTACCTGAAGCTGAGCTGTGTAACGGGTGTG 241  
|||  
QY 222 LeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGln 241  
|||  
Db 242 CTGGCCCTGCGCTGCGTGAAGATTCTCAGCACAGACAGCAAGCTCTCCTGGCCCCAG 301  
|||  
QY 242 LeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIle 261  
|||  
Db 302 CTCGGTTTGTCCAGGCTGTGGCGCTGCGCCAGGAGTCTTCCCTTCGGATC 361  
|||  
QY 262 ArgThrLeuProTrrPalaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPhe 281  
|||  
Db 362 CGCACCCCTCCCTGGGCTCATTTAAGCTGCCCTTCTTAACCTACTTTGGACTCTTT 421  
|||  
QY 282 GlnValHis 284  
|||  
Db 422 CAAGTTCAC 430

Search completed: May 8, 2003, 04:22:52  
Job time : 1338.02 secs



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:11:52 ; Search time 30.1297 Seconds  
(without alignments)  
1942.185 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAPFLTYGLQVH 284

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP virus:\*  
16: SP bacteriap:\*  
17: SP archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.0	325	5	Q9V5G2 drosophila
2	103	7.1	565	16	Q9KY66 streptomyc
3	101	7.0	643	16	Q9KZ17 streptomyc
4	100.5	7.0	375	16	Q9RR15 streptomyc
5	100	6.9	850	11	Q9J115 mus musculi
6	99.5	6.9	378	16	Q9AAB9 caulobacter
7	99	6.9	614	4	Q8WU14 homo sapien
8	99	6.9	855	4	Q9NKR9 homo sapien
9	99	6.9	878	4	Q9UFU7 homo sapien
10	98	6.8	206	16	Q9S2W5 streptomyc
11	97.5	6.8	1560	4	Q9FJP2 homo sapien
12	97	6.7	408	10	Q8S515 oryza sativ
13	97	6.7	926	4	Q9NYA0 homo sapien
14	97	6.7	1058	4	Q9Y4G2 homo sapien
15	96	6.6	629	10	Q8S1A6 oryza sativ
16	95.5	6.6	536	4	Q9HB96 homo sapien

17	95.5	6.6	655	16	Q9FBR7 streptomyc
18	94.5	6.5	748	5	Q8RTY0 trypanosoma
19	94.5	6.5	1696	11	Q9WTR8 rattus norv
20	94	6.5	937	16	Q93JD1 streptomyc
21	93.5	6.5	776	16	Q9RJ01 streptomyc
22	93.5	6.5	854	16	Q9F2P0 streptomyc
23	93.5	6.5	1038	10	Q9AS09 oryza sativ
24	93	6.4	507	2	Q9RMT5 streptomyc
25	93	6.4	1043	10	Q93VC8 oryza sativ
26	92.5	6.4	484	11	Q9CTA7 mus musculi
27	92.5	6.4	614	4	Q9NM41 homo sapien
28	92.5	6.4	707	11	Q925J8 mus musculi
29	92.5	6.4	1511	13	Q91BE2 fygu rubrip
30	92	6.4	351	12	Q937L3 orf virus
31	92	6.4	566	16	Q70005 streptomyc
32	92	6.4	845	4	Q96H68 homo sapien
33	92	6.4	977	4	Q8ET11 homo sapien
34	91.5	6.3	212	16	Q86312 mycobacteri
35	91.5	6.3	221	10	Q8RYZ6 oryza sativ
36	91.5	6.3	664	16	Q9HYV9 pseudomonas
37	91.5	6.3	3169	5	Q917V8 drosophila
38	91	6.3	681	16	Q9HUV6 pseudomonas
39	91	6.3	1217	4	Q60336 homo sapien
40	90.5	6.3	449	2	Q93HD5 streptomyc
41	90.5	6.3	760	11	Q8VD14 mus musculi
42	90.5	6.3	1100	2	Q937L3 deionococcus
43	90.5	6.3	1329	16	Q9CD30 mycobacteri
44	90	6.2	285	2	Q33629 saccharopol
45	90	6.2	422	16	Q9LDB9 streptomyc

## ALIGNMENTS

RESULT 1  
Q9V5G2 PRELIMINARY; PRT; 325 AA.  
AC Q9V5G2: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CG12919 protein.  
GN CG12919.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borstova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gilbert W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Houstoun K.A., Howland T.J., Mei M.H., Ibegwan C.,  
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Iel Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Port V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003831; AAF5848.1;  
 DR FlyBase; FBgn0034483; CG12919.  
 DR InterPro; IPR000478; TNE\_Family.  
 DR SMART; SM00207; TNE\_1.  
 DR PROSITE; PS00251; TNE\_1; 1.  
 DR PROSITE; PS50049; TNE\_2; 1.  
 SQ SEQUENCE 325 AA; 36862 MW; 655CCB69694F1A3A CRC64;

Query Match 8.0%; Score 116; DB 5; Length 325;  
 Best Local Similarity 25.4%; Pred. No. 0.07;  
 Matches 57; Conservative 32; Mismatches 97; Indels 38; Gaps 10;

OY 82 RASISAOEPAEELVAE---EDDPSELNPQTESODPAPFLN-----LVPR 127  
 DB 118 RKSINADYRNEONITQNHTELDEKSSNEATSKES--PALHHRHRRHRLHLVKRG 175  
 OY 128 RSAPGRKTRARAIAYEHVPRPGDGAQAGVGTVSGVEARINSSPLRYNRQGE 187  
 DB 176 ESLISARFEDSRP--AAHFHLSSRRRHGSM-GYHGDNYICNDNRNRYOG-HPOTRGV 231  
 OY 188 FIVRAGIYIXCOV-----HFDEKAVYLYKLDLVGVLAIRCLLEESATPAASLGFL 242  
 DB 232 LTVNTGTGLYYVAGIICYNNSHDQNFYF-----QGDTPRLQCLN---TVPTMPKRV 281  
 OY 243 RLCOVSGLLALPGSSLRIRTL--PMAHLKAAPLYEGLFV 283  
 DB 282 HTCHTSGLIHERNERIHLKDIHNDRAVLEGNKRSTFGIFKV 325

RESULT 2  
 O9KX66 PRELIMINARY; PRT; 565 AA.  
 AC O9KX66;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative peptidase.  
 GN SC04798 OR SCD63A.09C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1902;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Brown S.P., Harris D.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN-A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN-A3(2);  
 RX MEDLINE-97000351; PubMed-8843436;

RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
 RA Khashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).

RA [4]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL356832; CAB92661.1;  
 DR InterPro; IPR002886; Peptidase\_M37.  
 DR Pfam; PF01551; Peptidase\_M37; 1.  
 SQ SEQUENCE 565 AA; 58070 MW; 7D0418D480C6A284 CRC64;

Query Match 7.1%; Score 103; DB 16; Length 565;  
 Best Local Similarity 21.2%; Pred. No. 1.7;  
 Matches 70; Conservative 33; Mismatches 99; Indels 128; Gaps 16;

OY 2 SLDFEISA---RLPLRSTGSDGCAVROAPAPMARSRORRGCEPTALL-- 56  
 DB 290 ALLDQGEATPDARIPARA-GSRAGASRRRCP-----ALLTI 331  
 OY 57 -VPLALGL--GLALACGLLAVVSGSRASISAOEPAE----- 94  
 DB 332 AVPSACVSVAGIAASVSLTG--DESTETASAPDGNAPKPAANRKLDTQLT 389  
 OY 95 -LVAEED-----QDPSELNPQTESODPAPFLNRLVPRRSAPKGRTRARAIAY 146  
 DB 390 SLAAGADDFADRASRTGRIDLKAEDAEKRAQAEARERLPLKFAVKGSLAY 449  
 OY 147 EVHRRPGDG-----AAGVGTVSGVEARINSSPLRYNRQGE 187  
 DB 450 -----GQAGINMWSHTGIDPVLQGTVAADGVYR-----TOPNSAYGN 491  
 OY 188 FIVTRA--GLYLYICQVH-----FDEKAVYLYKLDLVGVLAIRCLLEESATPAAS 236  
 DB 492 MAITATADGTWTYCHLSYVPSGTIVKAGDAI-----ATSGPSGN 533  
 OY 237 SLGQRLCQVSGLLALPGSSLRIRTLPW 266  
 DB 534 STGPHLH-----FEVVPAGSSIDPLPW 556

RESULT 3  
 O9KZ17 PRELIMINARY; PRT; 643 AA.  
 AC O9KZ17;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical protein SC02220.  
 GN SC02220 OR SC1087.15.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1902;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Weizorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hapwood D.A.,  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL355752; CAB90868.1; .  
 KM Hypothetical protein.  
 SQ SEQUENCE 643 AA; 66537 MW; 1217A5C88E3CD419 CRC64;

Query Match 7.0%; Score 101; DB 16; Length 643;  
 Best Local Similarity 23.4%; Pred. No. 3;  
 Matches 68; Conservative 28; Mismatches 122; Indels 72; Gaps 10;

OY 12 RLP-----LPRSLGSRD-GGAVROAQ-----PPAPMAARRSRRGRGE 50  
 DB 165 RLPDGVROVLRAGVADPDALREADGVPGGYLLGSPEFPCSLQARPTDLRRRQH- 223  
 OY 51 PGTALLVPLALGLALACGLILAVNLSGSRASLSAOPPAQDELVAEDQPSSELNPT 110  
 DB 224 -----TKAALVAGALVCGALLGLPGDWCSPDGAAPPYAQNPAEALDPGRLTAA 277  
 OY 111 EESQDPAPFLNRLVPRRSGAPKGRKTRARAIAYEVHPRGDQGAQGVDTVSGMEE 170  
 DB 278 PAAMETSAKTDPSVMPARAGLTGDEELRLALA-----VWARGESVGSATGTATG--- 330  
 OY 171 ARINSSPLRYNRQIGEFIVTRAGLYLYYQVHFDEGRAVYKLDLVDGLALRLCEEF 230  
 DB 331 -----GP-----AGPPQLLYAGEVDTARVY-----VLHDLRLVRYRAERK 365  
 OY 231 SATASSL-----GPQLRLCOVSGLLALRPGSSLRRTTPMAHLKAA 272  
 DB 366 DGSAGALDEARTDAGRAATAVAVG-----RADGNVRYLTAPWYTKAA 411

## RESULT 4

O9RRH5 PRELIMINARY; PRT; 375 AA.  
 AC O9RRH5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein DR2516.  
 GN DR2516.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RI;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Usterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans RI.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE002081; AAF12062.1; .  
 DR TIGR: DR2516; .  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 375 AA; 39586 MW; 197B397658ED30B3 CRC64;

Query Match 7.0%; Score 100.5; DB 16; Length 375;  
 Best Local Similarity 24.0%; Pred. No. 1.7;  
 Matches 69; Conservative 27; Mismatches 84; Indels 107; Gaps 13;

OY 25 GAVROAQPAPMAA-----RSQRRRGRRGEGPT--ALLVPLALGLALACGLILAV 77  
 DB 5 GGHRCQPPRGAARRLRRGRAPRRGQPSHAPANORLRADGAALA-----LVA 58  
 OY 78 SIGSRASLSAOC--PAQDELVAEDQPSSELNPTQTESODPAPFLNRLVPRRSGAPKRT 136  
 DB 59 DVTNRPTPLAGSLPTLIDLAR-----QDLESDDSP----- 89  
 OY 137 RARRAIAYEVHPR-PGQDGAQAC--VDGTVSGMEARINSSPLRYNRQIGE-----F 188  
 DB 90 TDRLAVOARAAATFPRLPDPGAGFGHPISGTLGLE--RLSAASLRARHARFGORGYLVG 147  
 OY 189 IYTRAGLYLY-----CQVHFDEGRAVYKLDLVDGVALARCEERS 231  
 DB 148 VVADDAQAEVYELVAGLFAPDQPGEDRPMAPHFQELRLHL----- 188  
 OY 232 ATASSLGPQLRLCOVS-----GLALRPGSSLRRT 263  
 DB 189 ---PSATGEQTHLSLVAPGPDPDPWLPWQALATRLSGSASRLFT 232

## RESULT 5

O9JJI5 PRELIMINARY; PRT; 850 AA.  
 AC O9JJI5;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Mixed lineage kinase 3.  
 GN MK3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/OLA;  
 RX MEDLINE=20354997; PubMed=10894943;  
 RA Saridaki A., Ferraz C., Demaille J., Scherer G., Roux A.-F.;  
 RT "Genomic sequencing reveals the structure of the Kcnk4 and Map3k11  
 RT genes and their close vicinity to the Slp1 gene on mouse chromosome  
 RT 19.";  
 RL Cytogenet. Cell genet. 89:85-88(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTRAINS 1 SH3 DOMAIN.  
 DR EMBL: AF15142; AAF73281.1; .  
 DR HSSP: P29355; ISEM.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF000069; pkinase; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 DR ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;  
 KW ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;  
 KM Transferase.  
 SQ SEQUENCE 850 AA; 93199 MW; 8F026CB3532DC10E CRC64;

Query Match 6.9%; Score 100; DB 11; Length 850;  
 Best Local Similarity 29.2%; Pred. No. 5.1;  
 Matches 50; Conservative 10; Mismatches 73; Indels 38; Gaps 7;  
 OY 19 LGSRDGAVROAQPAPMAARRSQRNRG---RRGEGTALLVPLALGLALACGLILAA 75

```

Db 601 LGSSTPALNGNPPRSPPEEPERPRAGPTERNGSSGPKIOLALLRGTG-----LLA 654
QY 76 VVSG-----SR-----ASLSAQEPAP-----EELVAEEDDPSLNPQTEESODPA 117
Db 655 SLGIGRLQPPGGLSRESGESPTAPPPQMPSPCPPELPSPLIRLSOTYDANSSPTPG 714
QY 118 PELNRL-----VPRRSPAPKGRKTRARRAIAAHYEHPPPGDGAQAGVGT 164
Db 715 PLLDLGVSPGSAKSPRRETRGR-----TVSPPPGISRSAPGTPGT 758

RESULT 6
Q9AAB9 PRELIMINARY; PRT; 378 AA.
AC Q9AAB9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-MAR-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hlyd family secretion protein.
GN CC0683.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter.
OX NCBI_Taxid=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; Pubmed-11259647;
RA Nielsen M.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Potocha I., Heldelberg J.F., Alley M.R.K., Ohta N., Madock J.R.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Klonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utecherack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL PROC. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005743; AAK2268.1;
DR TIGR; CC0683;
DR InterPro; IPR003997; RtxD.
DR PRINTS; PR01490; RTXTOXIND.
KW Complete proteome.
SQ SEQUENCE 378 AA; 39659 MW; 88B818C437C15C8C CRC64;

Query Match 6.9%; Score 99.5; DB 16; Length 378;
Best Local Similarity 25.1%; Pred. No. 2.1;
Matches 73; Conservative 29; Mismatches 106; Indels 83; Gaps 13;

QY 19 LGSRDGAVRQAPPAAPMAARRSQRRGR---GEPGALLVPLALGLALACLGLILA 75
Db 36 LAAGGAGIVANVH-----REGDRVVGVPATATLTLSSALEGGDSFAVLSR--- 81
QY 76 VVSGSRASLSAQEPA---QEELVAEEDDPSLNPQTEESODPAFLNLRVPRRS--AP 131
Db 82 --SLGAGSSAAGSRAATQAALAAESRQ-----LIGRRALNRELAEARSLALQ 129
QY 132 KGRKTRAR-----RAIAAHYEHPPPGDGAQAGVGTGSGWEARINSSPLRYNQI 185
Db 130 TERLTLLARAEVARETTAAAGFLSPRE---LQARRSSELAVOOEASTLSQALSYEROI 185
QY 186 GEFTVTRAGLYLYCOVHFDEGKRAVYLKLD---LYDGVALARCLEESATNAASLSRQL 242
Db 186 GEV-----DARLAAPITDQAARAELASTLAGLEQ--QATQVEAGKRYV 227
QY 243 RLCOVSGILA-----LRPGSSLRITRLPFAHUKAPFL 275
Db 228 VVAIVAGRAVALPVEAGOTGVGAANVLTGDSALVAEL-YAISRAGFV 277

RESULT 7
Q8WU14 PRELIMINARY; PRT; 614 AA.
ID Q8WU14

```

```

AC Q8WU14; Q9BR73;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Histone deacetylase 7A.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 339-614 FROM N.A.
RC TISSUE-B-CELL;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020505; AAH20505.1;
DR EMBL; BC006453; AAH06453.1;
DR InterPro; IPR000286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 614 AA; 66186 MW; 2B638D8C866B502 CRC64;

Query Match 6.9%; Score 99; DB 4; Length 614;
Best Local Similarity 24.2%; Pred. No. 4.2;
Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

QY 11 RLPLPSLSRGDGAVR-----QAQPPAPMAAR-----RSQRRGR- 47
Db 63 RQISADLELDGEGPGQVDDGLEHRELGHGQPEARGPAPLQHPQVLLMEQRLRGL 122
QY 48 -RGEPGALLVPLALGLIACGLLAVVSLGSRASLSAQEPAQBELVAEEDDPSL 106
Db 123 PRGSGTDTVLPLAQGGHRPLS-----RAQSSPAAPASLSAPEASQARVLSSETPART 177
QY 107 NPQT-----EESQDP-----APFLNLRVPRRSAPKGRKTRAR 139
Db 178 LPFTTGLYDSVMLKHQCSGCDNSRHPHEHGRIGSIYSRLQERLRSQCCELRKASLE 237
QY 140 RAIAAHYEVH 149
Db 238 ELQSVHSERH 247

RESULT 8
Q9NRY9 PRELIMINARY; PRT; 855 AA.
ID Q9NRY9
AC Q9NRY9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Histone deacetylase 7.
GN HDAC7.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVICAL CARCINOMA;
RA Li S., Fische W., Verdin E., Walsh M.J.;
RT "A novel class II HDAC is associated with the transcriptional
RT homeodomain repressor CCAAT displacement protein.";
RL submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF239243; AAF63491.1;
DR InterPro; IPR000286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 855 AA; 92363 MW; 3157426AE1913DB6 CRC64;

```

Query Match 6.9%; Score 99; DB 4; Length 855;  
 Best Local Similarity 24.2%; Pred. No. 6.2;  
 Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

OY 11 RLPLPRSLGSRDGA---OAPPPAAR-----RSQRRGR- 47  
 DB 361 ROIPEAELETDGGPGGVVDGLEHRELHGQPEARGPAPQHPVLLMEQQLAGRL 420  
 OY 48 -RGEFTALVPLALGLALACGLLLAVVSLGSRASIAQEPAEDELVAEEDDPSRL 106  
 DB 421 PRGSTGDTVLLPLAOGGHRPLS-----RQSSPAPASISAEPPASQARVLSSSETPART 475  
 OY 107 NPOT-----EESODP-----APFLNRLVPRRSAPKGRKTRAR 139  
 DB 476 LPFTTGLIYDSVMLKHQSCSGDNHRHPEAGRIQSIWRLQGRGLRSGCECLGRKASLE 535  
 OY 140 RAIAAHYEVH 149  
 DB 536 ELQSVHSERH 545

## RESULT 9

O9UFU7 PRELIMINARY; PRT; 878 AA.  
 AC O9UFU7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 94.3 Kda protein (Fragment).  
 GN DKF2P586J0917.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=UTERUS;  
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL117455; CAB55935.1;  
 DR InterPro: IPR000286; His\_deacetylase.  
 DR Pfam: PF00850; Hist\_deacetyl; 1.  
 DR PRINTS: PRO1270; HDASUPER.  
 DR Hypothetical Protein.  
 FT NON TER 1  
 SO SEQUENCE 878 AA; 94330 MW; 26882E693960E10D CRC64;

Query Match 6.9%; Score 99; DB 4; Length 878;  
 Best Local Similarity 24.2%; Pred. No. 6.4;  
 Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

OY 11 RLPLPRSLGSRDGA---OAPPPAAR-----RSQRRGR- 47  
 DB 327 ROIPEAELETDGGPGGVVDGLEHRELHGQPEARGPAPQHPVLLMEQQLAGRL 386  
 OY 48 -RGEFTALVPLALGLALACGLLLAVVSLGSRASIAQEPAEDELVAEEDDPSRL 106  
 DB 387 PRGSTGDTVLLPLAOGGHRPLS-----RQSSPAPASISAEPPASQARVLSSSETPART 441  
 OY 107 NPOT-----EESODP-----APFLNRLVPRRSAPKGRKTRAR 139  
 DB 442 LPFTTGLIYDSVMLKHQSCSGDNHRHPEAGRIQSIWRLQGRGLRSGCECLGRKASLE 501  
 OY 140 RAIAAHYEVH 149  
 DB 502 ELQSVHSERH 511

RESULT 10  
 O9S2W5 PRELIMINARY; PRT; 206 AA.  
 AC O9S2W5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative membrane protein.  
 GN SC02091 OR SC4A10.24C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 NC NCI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Saunders D.C., Harris D.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351, PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 FT "A set of ordered cosmid and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)."  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL109663; CAB52000.1;  
 DR InterPro: IPR002965; P-richextensn.  
 DR PRINTS: PRO1217; PRICHEXTENS.  
 SO SEQUENCE 206 AA; 20839 MW; 90131DF9AAB059D7 CRC64;

Query Match 6.8%; Score 98; DB 16; Length 206;  
 Best Local Similarity 27.3%; Pred. No. 1.3;  
 Matches 48; Conservative 18; Mismatches 64; Indels 46; Gaps 7;

OY 20 GSRDGAVRQAQAPPPAARSRQRRGR-----RGEFTA--LLVPLALGL 63  
 DB 6 GPRDGGKI-QTPPRGGRVSRKPELKGRAARLARLLPTGSGGQAARAFVLLVVLGG 64  
 OY 64 GLALCIGLLAVVSLGSRASIAQEPAEDELVAEEDDPSRLNQTESQAPAFRLR 123  
 DB 65 GL-----IGLVV-----NSALSGSSRQLDDLKORTKELTDEQALORD 103  
 OY 124 VPRRSAPKGRKTRARRAIAAHYEVHPPGQAGVGVGTVSGWEARINSSPL 179  
 DB 104 I-DAYSARRALQRRAR-----ELGWPVGDPAPLDPDGTGVKGPSPAPAAATPL 151

RESULT 11  
 O96JP2 PRELIMINARY; PRT; 1560 AA.  
 AC O96JP2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE KIA1783 protein (Fragment).  
 GN KIA1783.

OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
 RT Prediction of the coding sequences of unidentified human genes. XX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.;  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF000612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 DR NON\_TER 1  
 PT SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

Query Match 6.8%; Score 97.5; DB 4; Length 1560;  
 Best local Similarity 25.5%; Pred. No. 17; Mismatches 117; Indels 67; Gaps 17;  
 Matches 75; Conservative 35;

QY 11 RLPLP-----RSLGSDGAVQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 ID 11 RLPLP-----RSLGSDGAVQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 AC 11 RLPLP-----RSLGSDGAVQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE Hypothetical 41.5 kDa protein.  
 BN OSJNB0061H20.4.  
 DS Oryza sativa (rice).  
 DS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 DC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 DC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=NIPPONBARE;  
 RA McComble W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,  
 RA Kuit K., Nascimento L., Zuberi T., Balla V., Bell M., Baker J.,  
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,

RESULT 12  
 DB 1158 KLINGAHSPPMLSPSPGKPPPAVAPRPKAPLQLGSSSIKKGQPLDLFG 1211  
 ID 1158 KLINGAHSPPMLSPSPGKPPPAVAPRPKAPLQLGSSSIKKGQPLDLFG 1211  
 AC 1158 KLINGAHSPPMLSPSPGKPPPAVAPRPKAPLQLGSSSIKKGQPLDLFG 1211  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE Hypothetical 41.5 kDa protein.  
 BN OSJNB0061H20.4.  
 DS Oryza sativa (rice).  
 DS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 DC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 DC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=NIPPONBARE;  
 RA McComble W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,  
 RA Kuit K., Nascimento L., Zuberi T., Balla V., Bell M., Baker J.,  
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,

RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;  
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
 RT OSJNB0061H20, from chromosome 10, complete sequence."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC113337; AM08840.1; -  
 DR Hypothetical protein.  
 KW SEQUENCE 408 AA; 41534 MW; CE4A6AB502840F45 CRC64;

Query Match 6.7%; Score 97; DB 10; Length 408;  
 Best local Similarity 26.5%; Pred. No. 3.8; Mismatches 84; Indels 66; Gaps 11;  
 Matches 60; Conservative 16;

QY 7 EISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 55  
 ID 7 EISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 55  
 AC 7 EISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 55  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP Hanski C., Gratchev A., Hartel-Schenk S., Ogorek D., Trendelenburg G.,  
 RA Hanski M.L., Hummel M., Riecken E.;  
 RT "Molecular cloning and characterisation of novel human protein with  
 RT two PH domains."  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ002220; CAB91652.1; -  
 DR InterPro: IPR000345; Cyt\_c\_heme\_bind.  
 DR InterPro: IPR002219; DAG\_PE\_bind.  
 DR InterPro: IPR001849; PH.  
 DR InterPro: IPR004012; Run.  
 DR Pfam: PF00169; PH; 1.  
 DR Pfam: PF02759; RUN; 1.  
 DR SMART: SM00109; C1; 1.  
 DR SMART: SM00233; PH; 2.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE: PS50081; DAG\_PE\_BIND\_DOM\_2; 1.  
 DR PROSITE: PS50003; PH\_DOMAIN; 1.  
 DR SEQUENCE 926 AA; 103198 MW; 66B2618C12C3B27B CRC64;

Query Match 6.7%; Score 97; DB 4; Length 926;  
 Best local Similarity 24.2%; Pred. No. 10; Mismatches 125; Indels 56; Gaps 13;  
 Matches 69; Conservative 35;

QY 4 LDEISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 63  
 ID 4 LDEISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 63  
 AC 4 LDEISARRLPRLPSLGRDGAQVQAQPPA---ARRSQRRC--RGEPTALL 63  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE Hypothetical 41.5 kDa protein.  
 BN OSJNB0061H20.4.  
 DS Oryza sativa (rice).  
 DS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 DC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 DC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=NIPPONBARE;  
 RA McComble W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,  
 RA Kuit K., Nascimento L., Zuberi T., Balla V., Bell M., Baker J.,  
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:47 ; Search time 10.6012 Seconds

(without alignments)  
1111.129 million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPRLSLG.....PMAHLKAPLFTYGLRQVH 284

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	87.8	249	TN12_HUMAN	O43508 homo sapien
2	1020	70.6	225	TN12_MOUSE	O54907 mus musculu
3	109.5	7.6	272	TNF5_CHICK	O91848 gallus gall
4	108.5	7.5	260	TNF5_CANFA	O91626 canis famli
5	106.5	7.4	254	TNF9_HUMAN	P41273 homo sapien
6	97	6.7	441	CG22_ANTNA	P34801 antitritinu
7	95.5	6.6	201	TNF8_MOUSE	O94146 macropus eu
8	94	6.5	280	TNF6_MACMU	P52154 macococcus
9	93.5	6.5	690	RHO_MICLU	O11162 mycobacteri
10	93	6.4	310	TN14_HUMAN	O43557 homo sapien
11	92	6.4	240	TN14_MOUSE	O93041 cercocebus
12	92	6.4	280	TNF6_CERTO	O93041 cercocebus
13	92	6.4	902	NEC4_HUMAN	O14934 homo sapien
14	90.5	6.3	760	MLH1_MOUSE	O91k91 mus musculu
15	90.5	6.3	814	CADFI_HUMAN	P55291 homo sapien
16	90	6.2	707	JIP1_MOUSE	O94w19 mus musculu
17	89	6.2	280	MDCB_KLEPN	P71422 klebsiella
18	89	6.2	316	TN11_MOUSE	O35235 m tumor nec
19	88.5	6.1	278	TNF6_RAT	P36940 rattus norv
20	88.5	6.1	281	TNF6_HUMAN	P48023 homo sapien
21	87	6.0	197	TNF8_RABIT	P10154 oryctolagus
22	87	6.0	204	TNF8_BOVIN	O06600 bos taurus
23	87	6.0	291	TN10_MOUSE	P50592 mus musculu
24	86.5	6.0	250	TNF6_MOUSE	O94x14 macropus eu
25	86	6.0	139	TOFB_BACSU	P34467 bacillus su
26	86	6.0	205	TNFB_MARMO	O91m09 marmota mon
27	86	6.0	241	TN13_MOUSE	O04477 mus musculu
28	86	6.0	933	VGIB_HSVAI	O04466 herpesvirus
29	85.5	5.9	928	VGIB_HSVBP	P12471 bovine hepi
30	85.5	5.9	932	VGIB_HSVBC	P12640 bovine hepi
31	85	5.9	372	LMXB_MOUSE	O88609 mus musculu
32	85	5.9	379	LMXB_HUMAN	O60663 homo sapien
33	84.5	5.9	401	AROC_MYCTU	P95013 mycobacteri

34	84	5.8	284	1	HXL1_HUMAN	O43763 homo sapien
35	84	5.8	310	1	TNFC_MARMO	O91m10 marmota mon
36	84	5.8	575	1	MIS_PIG	P79295 sus scrofa
37	84	5.8	825	1	ICP0_HSV2H	P28284 herpes stimp
38	83.5	5.8	416	1	RAGE_BOVIN	O28173 bos taurus
39	83.5	5.8	505	1	TUB_MOUSE	P50586 mus musculu
40	83.5	5.8	545	1	RTN2_HUMAN	O75298 homo sapien
41	83.5	5.8	885	1	VGIB_HSV2S	P24994 herpes stimp
42	83	5.7	228	1	B10D_PSEAE	O91614 pseudomonas
43	83	5.7	653	1	APF1_MOUSE	O03157 mus musculu
44	83	5.7	936	1	FHL1_YEAST	P39521 saccharomyc
45	83	5.7	1217	1	AF4_MOUSE	O88573 mus musculu

## ALIGNMENTS

RESULT 1	ID	TN12_HUMAN	STANDARD:	PRT:	249 AA.
AC	O43508	OBMU27;			
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).				
DE	TNFSF12 OR APO3L OR DR3LG.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.				
RC	TISSUE-Tonsil, and Fetal liver;				
FX	MEDLINE=98070415; PubMed=9405449;				
RA	Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;				
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."				
RL	J. Biol. Chem. 272:32401-32410(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Fetal kidney;				
FX	MEDLINE=9828355; PubMed=9560343;				
RA	Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A., Ashkenazi A.;				
RT	"Identification of a ligand for the death-domain-containing receptor APO3."				
RL	Curr. Biol. 8:525-528(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Tonsil;				
RA	Strauberg R.;				
RL	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.				
RN	[4]				
RP	FUNCTION.				
FX	PubMed=10085077;				
RA	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;				
RT	"TWEAK induces angiogenesis and proliferation of endothelial cells."				
RL	J. Biol. Chem. 274:8455-8459(1999).				
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.				
CC	-1- SUBUNIT: HOMOTRIMER (POTENTIAL).				
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.				
CC	-1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas, skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peripheral blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.				
CC	-1- PTM: The soluble form derives from the membrane form				

by proteolytic processing.  
 -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 -1- CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 125.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL: AF030099; AAC51923.1; -  
 EMBL: AF05872; AAC39724.1; -  
 EMBL: BC019047; AAH19047.1; ALT\_FRAME.  
 Genew: HGNC:11927; TNFSF12.  
 MIM: 602695; -  
 InterPro: IPR000478; TNF\_family.  
 Pfam: PF00229; TNF; 1.  
 SMART: SM00207; TNF; 1.  
 PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 PROSITE: PS50049; TNF\_2; 1.  
 Cytokine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CHAIN 1 249  
 CHAIN 94 249  
 DOMAIN 1 21  
 TRANSMEM 22 42  
 DOMAIN 43 249  
 SITE 93 94  
 CARBOHYD 139 139  
 SEQUENCE 249 AA: 27216 MW: E60843361C28B8A CRC64;

Query Match 87.8%; Score 1268; DB 1; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 2e-93;  
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

36 MAARRSGRRGRGEPCTALLVPLALGLALCLGLLAVLSGSNASISAQEPAGEEL 95  
 1 MAARRSGRRGRGEPCTALLVPLALGLALCLGLLAVLSGSNASISAQEPAGEEL 60  
 96 VAEDDQPSSELNPQTEESQDPAPFLNLVLRPRRSAPRKTRARRAIAHYEYHPRGQD 155  
 61 VAEDDQPSSELNPQTEESQDPAPFLNLVLRPRRSAPRKTRARRAIAHYEYHPRGQD 120  
 156 GAAGAGVGTYSWGWEARINSSPLRYNRQIGEFIVTRAGLYLYCOVHDEGRAVYLKLD 215  
 121 GAAGAGVGTYSWGWEARINSSPLRYNRQIGEFIVTRAGLYLYCOVHDEGRAVYLKLD 180  
 216 LVNDGVYALRCLEEFSTAASSIGPOLRLCOVSGILALRPSSLRITLPMWHLKAAPFL 275  
 181 LVNDGVYALRCLEEFSTAASSIGPOLRLCOVSGILALRPSSLRITLPMWHLKAAPFL 240  
 276 TYFGLEFOVH 284  
 241 TYFGLEFOVH 249

RESULT 2  
 ID TN12\_MOUSE STANDARD: PRT: 225 AA.  
 AC 054907: 09CMP2:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (Fragment).  
 GN TNFSF12.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Peritoneal macrophage;  
 RX MEDLINE=98070415; PubMed=9405449;  
 RA Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,  
 RA Hesson C., Garcia I., Browning J.L.;  
 RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis";  
 RL J. Biol. Chem. 272:32401-32410(1997).  
 RN [2]  
 RP SEQUENCE OF 83-225 FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Retina;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gasterland T., Glasl C., King B., Kochiba H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirni L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: Binds to TNF14 and possibly also to TNFSF12/Apo3. Weak inducer of apoptosis in some cell types. Promotes angiogenesis and the proliferation of endothelial cells. Mediates NF-kappaB activation (By similarity).  
 CC -1- SUBUNIT: Homotrimer (Potential).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- PTM: The soluble form is produced from the membrane form by proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL: AF030100; AAC53517.1; -  
 EMBL: AK020909; BAB32249.1; -  
 MGD: MGI:1196259; Tnf5f12.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 FT NON-TER 1  
 FT CHAIN <1 225  
 FT CHAIN 70 225  
 FT TRANSMEM <1 21  
 FT DOMAIN 22 225  
 FT SITE 69 70

EXTRACELLULAR (POTENTIAL).  
 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 CLEAVAGE (BY SIMILARITY).

FT	DISULFID	167	186	POTENTIAL.
SO <td>CARBONHD</td> <td>115</td> <td>115</td> <td>N-LINKED (GLCNAC...) (POTENTIAL).</td>	CARBONHD	115	115	N-LINKED (GLCNAC...) (POTENTIAL).
SO <td>SEQUENCE</td> <td>225 AA;</td> <td>24/81 MW;</td> <td>90CA12CC0480659B CRC64;</td>	SEQUENCE	225 AA;	24/81 MW;	90CA12CC0480659B CRC64;
	Query Match	70.6%	Score 1020;	DB 1; Length 225;
	Best Local Similarity	88.8%;	Pred. No. 7.4e-74;	
	Matches 199;	Conservative 9;	Mismatches 16;	Indels 0; Gaps 0;
QY	1	LGGLIATLACGLILLAVVSLIGSRASISAOEPAOEELVAEEDDPSELNPOTESODPAEFL	120	
DB	2	LSLGIATLACGLILLAVVSLIGSWATLSAQEPSEELTAEDRRPELPANPOTESODVPEFL	61	
QY	121	NRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDGYISGWEARINSSPLR	180	
DB	62	EOLVPRRSAPKGRKARRRRAIAAHYEVHPRPGDGAQAGVDGYISGWEERKINSSPLR	121	
QY	181	YNROIGETVIRAGIYYLYCOVHPDEGKAVYLKLDLVGVYALCLSEFSATASSIGP	240	
DB	122	YDRQIGETTVIRAGIYYLYCOVHPDEGKAVYLKLDLVGVYALCLSEFSATASSIGP	181	
QY	241	QLRLQVSGLLALRPSSSLIRITLPMWAKKAPFLTYEGLRQVH	284	
DB	182	QLRLQVSGLLALRPSSSLIRITLPMWAKKAPFLTYEGLRQVH	225	
RESULT 3				
TFN5_CHICK	TFN5_CHICK	STANDARD;	PRT;	272 AA.
ID	Q918D8:			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-			
DE	L) (C0154 protein).			
GN	TFN5F5 OR CD40LG OR CD40L.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosaustra; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=White leghorn; TISSUE=spleen;			
RA	Tregaskes C.A., Young J.R., Burnside J.;			
RL	Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell			
CC	proliferation in the absence of IL-4. Involved in immunoglobulin			
CC	class switching (By similarity).			
CC	-1- SUBUNIT: Homotrimer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an			
CC	extracellular soluble form (By similarity).			
CC	-1- PM: The soluble form derives from the membrane form by			
CC	proteolytic processing (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: AJ243435; CAB95748.1; -			
DR	HSSP: P29965; IALY.			
DR	InterPro: IPR003263; TNF_5.			
DR	InterPro: IPR003636; TNF_abc.			
DR	InterPro: IPR000478; TNF_family.			
DR	Pfam: PF00229; TNF_1.			
DR	ProDom: PD002012; TNF_abc; 1.			
DR	ProDom: PD008600; TNF_5; 1.			
DR	SMART: SM00207; TNF_1.			

DR PROSITE; PS00251; TNE\_1; 1.  
RW PROSITE; PS00049; TNE\_2; 1.  
KW Cytochrome; Transmembrane; Glycoprotein; Signal-anchor.  
FT CHAIN 1 272  
TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY MEMBER 5, MEMBRANE FORM.

FT CHAIN 111 272  
TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY MEMBER 5, SOLUBLE FORM (BY SIMILARITY).  
CYTOPLASMIC (POTENTIAL).  
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
EXTRACELLULAR (POTENTIAL).  
CLEAVAGE (BY SIMILARITY).  
POTENTIAL.

FT SITE 45 272  
DISULFID 110 111  
CARBOHYD 190 229  
CARBOHYD 124 124  
CARBOHYD 146 146  
CARBOHYD 251 251  
SEQUENCE 272 AA; 30862 MM; 5409P24ABE53CCD7 CMC64;

Query Match Best Local Similarity 7.6%; Score 109.5; DB 1; Length 272;  
Matches 42; Conservative 26; Mismatches 58; Indels 33; Gaps 7;

OY 139 RRAIAHVEVPRPGQDAAGVDSG--EEAIISSSPLRVNRGGEIYRAGLY 197  
Db 133 KOPITTHLA-----GVKSTTVRYALKMTTSAPSLISTYE--GKLKEVAAGLY 182

OY 198 LYCVHFPEGRKA-----YLKLDLLVDGYALRCIEEFSAATASSLGQLRLCY-- 247  
Db 183 IYSQSFCTKAAPAFLPIYLYLPMEDRL-NKGDIHTSYTA-----LCEDIQS 233

OY 248 ---SGLLALRPSSLRIRTLPAHLKAAPFLTYGLFGV 283  
Db 234 IREGVFELRGQDMVFVNVTDTAAVNPVNTYFGMFRL 272

RESULT 4  
ID TNFS\_CANFA STANDARD; PRT; 260 AA.  
AC O97626;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).  
GN TNFSF5 OR CD40LG OR CD40L.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_Taxid=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hosie M.H., Willett B.J.;  
RT "Advant properties of canine CD40L."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.

-I- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the absence of co-stimulus as well as IgE production in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).  
-I- SUBUNIT: HOMODIMER (By similarity).  
-I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM (By similarity).  
-I- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).  
-I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF086711; AAD04375.1; -

```

DR HSP: P29965; ITALY.
DR InterPro: IPR003263; TNF_5.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR ProDom: PD008600; TNF_5; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS0049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 260
FT FT
FT CHAIN 112 260
FT FT
FT DOMAIN 1 22
FT TRANSMEM 23 46
FT FT
FT DOMAIN 47 260
FT SITE 111 112
FT DISULFID 177 217
FT CARBOHYD 239 239
FT SEQUENCE 260 AA; 28688 MW; 604F69A19E98EB70 CRC64;

Query Match
Best Local Similarity 25.5%; Score 108.5; DB 1; Length 260;
Matches 53; Conservative 23; Mismatches 69; Indels 63; Gaps 8;

QY 89 EPAQELVAEE--DDPPS-ELNPTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAH 145
DB 103 EKKKEENIMQKDDPPRIAAHVISEASSNPASVL-----RMARKGYYTISNLSVL- 154
QY 146 YEVHPRPGDGAQAGVDGTSCMEARINSSPLRYNRQIGFEIYTRAGLYLYCOVHPD 205
DB 155 -----ENGRKQ-----LAKRGSLYVYVQVTF 177
QY 206 ECKAYVYKLDLVYDGLALRCEFSAT-----AASSIGPOLRLCOVS---GLLALRP 255
DB 178 SNRASSQAPF-----VASILCHSPSGTERVLLRAASSSGSKPCGQGSIHGVEFLHP 232
QY 236 GSLRIRFTLPMVHLKAPFLTYEGLFOY 283
DB 233 GASVAVNTDPSQVSHGTGFTSFGLLKL 260

RESULT 5
TNF9_HUMAN STANDARD: PRT: 254 AA.
AC P41273:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 9 (4-1BB ligand) (4-1BB).
GN TNFSF9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=94374434; Pubmed=8088337;
RA Alderson M.R., Smith C.A., Tough T.W., Davis-Smith T., Armitage R.J.,
RA Falk B., Roux E., Baker E., Sutherland G.R., Din W.S., Goodwin R.G.,
RT "Molecular and biological characterization of human 4-1BB and its
RT ligand."
RL Eur. J. Immunol. 24:2219-2227(1994).
CC -1- FUNCTION: Cytokine that binds to TNFRSF9. Induces the
CC proliferation of activated peripheral blood T cells. May have a
CC role in activation-induced cell death (AICD). May play a role in
CC cognate interactions between T cells and B cells/macrophages.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PLACENTA, LUNG, SKELETAL

```

```

CC MUSCLE AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL: U03398; AAA53134.1;
DR Genbank: HGNC:11939; TNFSF9.
DR MIM: 606182;
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS0049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphism.
FT CHAIN 1 28
FT TRANSMEM 29 49
FT FT
FT DOMAIN 50 254
FT FT
FT DOMAIN 35 41
FT VARIANT 17 17
FT SEQUENCE 254 AA; 26624 MW; 82751F34563E508 CRC64;

Query Match
Best Local Similarity 26.5%; Score 106.5; DB 1; Length 254;
Matches 74; Conservative 31; Mismatches 93; Indels 81; Gaps 13;

QY 32 PPAPAAARRSQRRGRGCEPTALLVPLAIGLALACIGLLAVSL-GSPAST-SQOE 89
DB 16 PPAP-----RARCRRVLP-WALVAGLLILLLLAAACVFLACRWANSGAASGSAAAS 67
QY 90 PAQOEIVAEEDDPSELNPTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAHYEH 149
DB 68 PRLRE-----GPESLP-----DDPAGLIDL-----RQGMFAQLVQNVLL- 102
QY 150 PRGGDGAQAGVDGTSCGME--ARINSSPLRYNRQIGFEIYTRAGLYLYCOVHPD 206
DB 103 -----IDGPLSTSDPGLAGVSLTGLSTYKEDTKELVAKAVYYVFFQ----- 146
QY 207 GRAYVYKLDLVDS-----VLALRCEFSATASSIGPOLRLCOVS----- 249
DB 147 -----LELRVYVAGSGSVSLAHLQPLRSGAAGALALVYDLPASSEARNSAFRG 201
QY 250 -LLALRPSSSLRIRFTLPMVHLKAPFL-----TYEGLFOY 283
DB 202 RLHLISAGQRLGVHLTEARARHAMQLTGATVGLFPRV 240

RESULT 6
G222_ANTMA STANDARD: PRT: 441 AA.
ID CG222_ANTMA
AC P34801;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G2/mitotic-specific cyclin 2.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.
ON NCBI_TaxID=4151;
RX MEDLINE=94148008; Pubmed=8313906;
RA Robert P.R., Coen E.S., Murphy G.J.P., Doonan J.H.;
RT "Patterns of cell division revealed by transcriptional regulation of
RT genes during the cell cycle in plants."

```

```

RL      EMBO J.13:616-624(1994).
CC      -I- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M
CC      (MITOSIS) TRANSITION. G2/M CYCLINS ACCUMULATE STEADILY DURING G2
CC      AND ARE ABRUPTLY DESTROYED AT MITOSIS.
CC      -I- SUBUNIT: INTERACTS WITH THE CDC2 AND CK2 PROTEIN KINASES TO FORM
CC      A SERINE/THREONINE KINASE Holoenzyme COMPLEX. THE CYCLIN SUBUNIT
CC      IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC      -I- DEVELOPMENTAL STAGE: ACCUMULATES STEADILY DURING G2 AND IS
CC      ABRUPTLY DESTROYED AT MITOSIS.
CC      -I- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC      -----
CC      THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC      THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON
CC      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC      ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-sib.ch/announce/).
CC      OR SEND AN EMAIL TO license@isb-sib.ch).
CC      -----
CC      EMBL; X76123; CAAS3729.1; -.
CC      PIR; S41710; S41710.
CC      HSSP; P30274; IVIN.
CC      INTERPRO; IPR004366; Cyclin.
CC      INTERPRO; IPR004367; Cyclin_Cterm.
CC      Pfam; PF00134; cyclin; 1.
CC      Pfam; PF02984; cyclin_C; 1.
CC      SMART; SM00385; CYCLIN; 2.
CC      PROSITE; PS00292; CYCLINS; 1.
CC      CYCLIN; Cell cycle; Cell division; Mitosis.
CC      KW      SEQUENCE 441 AA; 49205 MW; EED4C037C98880A7 CRC64;

```

Query Match	6.7%	Score 97	DB 1	Length 441
Best Local Similarity	24.0%	Pred. No. 1.6		
Matches 70	Conservative 44	Mismatches 104	Indels 74	Gaps 15
QY	25	GAVRQAGPAPMAARRSOR	-----RGRGEPGTALLVPLATGLATLACGLL	73
DB	19	GAIKOKS---	MAVEKKRRFALGDIGNVTVGVGEGKLLPQVSRPRTIRG	69
QY	74	-----LAVYSIGSRASLS	-----AQEPAQEEIYAEDDOPSELTPOTEEQ	114
DB	70	ANAEAAAEEENKNSLAVNNAKGDAGLPIKRAVARPVQKTKSKPQETILETISPTKEKK	-----	129
QY	115	DEAPELNLRVPRRS---	APKGRKTRARAIYAAHYEHPREGQ-----DGAQAGVDGTV	165
DB	130	--APVLEKEITGEKSLKKKAPTLTSTLVARSKAAV-	VRTKREQIDVIDADAVNNDLAV	186
QY	166	SCWEE-----ARINSSPLRY---	NQIGCFYTRAGLYLYTCQYHFD---ECKAYYL	212
DB	187	VEYVEDMYKFKYKSAENDSRPHDYMSOPEINCKM--	RAILLDMLVQVHHKPELSPEITYL	244
QY	213	KIDLVDVGLALRCL---	LEEFSAVAASSLGFOLRLCOVSGGLALRPGS	257
DB	245	TIN-IVDRIYLSKTSRRELQLLGNSMILASKYEINAPENVNDLYCISDGS	295	
RESULT 7				
TNFB_MACEU	TNFB_MACEU	STANDARD:	PRT:	201 AA.
AC	09XT48:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lymphotoxin- $\alpha$ precursor (LT- $\alpha$ ) (TNF- $\beta$ ) (Tumor necrosis factor ligand superfamily member 1).			
DE	LTA OR TNESF1 OR TNFB.			
OS	Macroporus eugeni (Tammara wallaby).			
OC	Eukaryota; Metazoa; Chordata; Ctenata; Vertebrata; Euteleostomi;			
OC	Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.			
OX	NCBI_TaxID=9315;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20284948; PubMed=10826697;			

```

RA Harrison G.A., Deane E.M.:
RT "cDNA cloning of lymphotoxin alpha (LT-alpha) from a marsupial,
RT Macropus eugenii."
RL DNA Seq. 10:399-403(2000).
CC
CC -1- FUNCTION: Cytokine that in its homotrimeric form binds to
CC TNFRSF1A/TNFR1, TNFRSF1B/TNFR and TNFRSF14/HVEM. In its
CC heterotrimeric form with LTB binds to TNFRSF3/LTRR. Lymphotoxin is
CC produced by lymphocytes and cytotoxic for a wide range of tumor
CC cells in vitro and in vivo.
CC
CC -1- SUBUNIT: Homotrimer, and heterotrimer of either two LTB and one
CC LTB subunits or (less prevalent) two LTA and one LTB subunits (By
CC similarity).
CC
CC -1- SUBCELLULAR LOCATION: Secreted (homotrimer) and type II membrane
CC protein (heterotrimers) (by similarity).
CC
CC -----
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC	EMBL; AF119336; AAD41773.1; ..
DR	HSSP; P01374; 1TNR.
DR	InterPro; IPR003636; TNF_abc.
DR	InterPro; IPR00478; TNF_family.
DR	Pfam; PF00229; TNF; 1.
DR	PRINTS; PR01234; TNECROSISRCT.
DR	PRODOM; PD002012; TNF_abc; 1.
DR	SMART; SM00207; TNF; 1.
DR	PROSITE; PS00251; TNF_1; 1.
DR	PROSITE; PS50049; TNF_2; 1.
KW	Cytokine; Glycoprotein; Cytotoxin; Signal.
FT	SIGNAL
FT	CHAIN
FT	CARBOHD
SO	SEQUENCE
	Query Match
	Best Local Similarity 23.2%; Score 95.5; DB 1; Length 201;
	Matches 44; Conservative 27; Mismatches 84; Indels 35; Gaps 9;
OY	107 NQTESQDPAPFLRLVPRRSARKGRTRARRAI--AAHEVHPRGQDCAQAGVDGT 164
DB	30 NEDNSHSSSPAE-----PQTAOHLGSKSLRETLKPAAH-----VEDPSVDSIH-- 75
OY	165 VSGMEEARINSSSP-LRYNROI--GEFIYTRAGLYLYCOVFDEGKA-----VYLKL 214
DB	76 ---W---RANTHAALRHREFSLSNNSLVPYTGILFYVSQVVFSGASCSCEITPLTYLSH 129
OY	215 DLDVG--VIALRCLEESATPAASSLGQLRLCOVSGILLRPSSLIRIPLPAHHKA 271
DB	130 EVLLPSSKYQHVPPLLSAQKSVCSGYQPMMSYVGAVALLTQGDRSLTYDGVSILLQ 189
OY	272 APFLTFPGLF 281
DB	190 SPSSVFFGAF 199
	RESULT 8
	TNF6_MACMU
ID	TNF6_MACMU STANDARD: PRT: 280 AA.
AC	Q9MYL6; Q9BDM5;
DT	15-JUN-2002 (Rel. 41, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Tumor necrosis factor ligand superfamily member 6 (Fas antigen ligand) (CD95L protein).
GN	TNFSF6 OR FASL OR CD95L.
OS	Mus musculus (house mouse), Macaca mulatta (rhesus macaque), Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and Macaca nemestrina (pig-tailed macaque).

CC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Ceroptreidae;
OC	Cercopithecoidea; Macaca.
OX	NCBI_TaxID=9544, 9541, 9545;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	SPECIES=M.mulatta; TISSUE=Lymphocytes;
RX	MEDLINE=2138618; PubMed=1149153;
RA	Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA	Weiss W.R., Ansari A.A.;
RT	"Cloning, sequencing, and homology analysis of nonhuman primate
RT	Fas/Fas-ligand and co-stimulatory molecules.";
RL	Immunogenetics 53:315-328(2001).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	SPECIES=M.mulatta, M.fascicularis, and M.nemestrina;
RA	Kiril' Y., Inoue T., Yoshino K.;
RL	Submitted (NOV-1999) to the EMBL/Genbank/DBD databases.
CC	-1- FUNCTION: Cytochrome that binds to TNFRSF6/FAS, a receptor that
CC	transduces the apoptotic signal into cells. May be involved in
CC	cytotoxic T cell mediated apoptosis and in T cell development.
CC	TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC	peripheral tolerance, in the antigen-stimulated suicide of mature
CC	T cells, or both. Binding to the decoy receptor TNRSF5B/DCR3
CC	modulates its effects (By similarity).
CC	-1- SUBUNIT: Homotrimer (potential).
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC	similarity).
CC	-1- PM: The soluble form derives from the membrane form by
CC	proteolytic processing (By similarity).
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on keys
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC	or send an email to license@sdb.sdb.ch).
CC	-----
DR	EMBL; AF344856; AAK37539.1; -
DR	EMBL; AB035138; BAA90294.1; -
DR	EMBL; AB035139; BAA90295.1; -
DR	EMBL; AB035140; BAA90296.1; -
DR	HSSP; P01375; ATSV.
DR	InterPro: IPR003636; TNF_abc.
DR	InterPro: IPR000478; TNF_family.
DR	Pfam: PF00229; TNF; 1.
DR	PRINTS: PR01334; TNCRSISRCT.
DR	PRODOM: PD002012; TNF_abc; 1.
DR	SMART; SM00207; TNF; 1.
DR	PROSITE; PS00251; TNF_1; 1.
DR	PROSITE; PS50049; TNF_2; 1.
KW	Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT	CHAIN 1 280
FT	CHAIN 129 280
FT	DOMAIN 1 80
FT	TRANSMEM 81 101
FT	DOMAIN 102 280
FT	DOMAIN 4 69
FT	DOMAIN 4 64
FT	SITE 128 129
FT	DISULFD 201 232
FT	CARBOHYD 183 183
FT	CARBOHYD 249 249
FT	CARBOHYD 259 259
FT	CONFLICT 60 60
FT	SEQUENCE 280 AA; 31367 MW; PDB264D61A132EB4 CNC64;
Query Match	6.5%; Score 94; DB 1; Length 280;

```

Best Local Similarity 20.8%, Pred. No.1:6;
Matches 62; Conservative 45; Mismatches 117; Indels 74; Gaps 14;

QY 13 LPLPRSLGSRDGCAGVRAQAPPAAMARRSRQRGRGEPGALLVPLAL----- 61
Db 30 LPCPTSVRRRQGRPRPPPPPPPLP-----PPPSPLPPLPLPKKRGNHST 78
QY 62 GLGLAALAGLLLVAVSLGSRASLSAQRPADDELVAEDDQPSSELPQTESQDAPFLN 121
Db 79 GLCLLVMEFMYLVAVLGIG--LGMEFLHKLKEL-----AELRESTOKHATSLSEK 128
QY 122 RLVRPRRSAPKGRGTBARRAIAAHVEYVPRRGDQAGVDTGVSAGWEA-RINSSPLR 180
Db 129 QIGHP---SPPEKKEQK--VAHLTKPRNSRMPLE-----WEDTGYILSSVK 174
QY 181 YNRQIGEEIVTRAGLYLYLCVHDEGKA-----VYLKLD-----LLVDGVALR 225
Db 175 YKK--GGLVINETGYLFYYSKYVF--RGSCGNLPLSHKVYRNRSKYPQDVLWMECKMMSY 231
QY 226 CLEESATAASSLCGPNLFCOVSGLLALPSSSLRIFRLPNAHLKAAPFLYFGIFQV 283
Db 232 CTGGGMASHSYLGAVENTLSADHLY-----VAVSELSLVNEEESQ--TEFGLYKL 280

RESULT 9
RHO_MTCU STANDARD; PRT: 690 AA.
AC P52154;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription termination factor rho.
GN RHO.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Actinobacteria; Actinobacteris (class); Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
RC STRAIN-EM;
RX MEDLINE=96132802; PubMed=8557681;
RA Nowatzke W.L., Richardson J.P.;
RT "Characterization of an unusual Rho factor from the high G + C gram-
RL positive bacterium Micrococcus luteus.";
RN J. Biol. Chem. 271:742-747(1996).
RN [2]
RP SEQUENCE OF 205-690 FROM N.A.
RC STRAIN-EM;
RX MEDLINE=94327472; PubMed=8051015;
RA Opperman T., Richardson J.P.;
RT "Phylogenetic analysis of sequences from diverse bacteria with
RL homology to the Escherichia coli rho gene.";
RN J. Bacteriol. 176:5033-5043(1994).
RN [3]
RP REVISION TO 500.
RC STRAIN-EM;
RA Nowatzke W.L.;
RL Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM
CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S
CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE
CC DNA TEMPLATE. RNA-DEPENDENT NTPASE WHICH UTILIZES ALL FOUR
CC RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DMP AS SUBSTRATES, BUT
CC HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC EMBL: L27277; AAB18671.1; -

DR HSSP; P03002; 1A63.

DR InterPro; IPR000194; ATPase\_a/bcentre.

DR InterPro; IPR002059; Cold\_shock.

DR InterPro; IPR004665; Term\_rho.

DR Pfam; PF00006; ATP-synt\_ab; 1.

DR SMART; SM00357; CSP; 1.

DR TIGRfams; TIGR00767; rho; 1.

KW Transcription termination; Helicase; ATP-binding; RNA-binding.

FT INIT\_MET 0

FT DOMAIN 30 35 RNA-BINDING (RNP2) (BY SIMILARITY).

FT DOMAIN 318 321 RNA-BINDING (RNP1) (BY SIMILARITY).

FT NP\_BIND 441 448 ATP (POTENTIAL) (BY SIMILARITY).

FT CONFLICT 291 291 G->P (IN REF. ?; AA SEQUENCE).

SO SEQUENCE 630 AA; 75030 MW; F77C4C75EE1B8958 CnC64;

Query Match

Best Local Similarity 25.9%; Score 93.5; DB 1; Length 690;

Matches 43; Conservative 19; Mismatches 63; Indels 41; Gaps 7;

QY 2 SLPLPESARLPPLPRISGSDGCAVROAPPAPMAARRSGRRGEPGTALLVPLAL 61

DB 55 SVADRDAAERAAQAPAPAAETAPAAASSEDAP-AAERPARRRSRADADTS--APAA 111

QY 62 GLGLALACIGLLAVSLGSRASLSAOPAEELVAEDODPSSELNPQTESODAPPLN 121

DB 112 QDQQA-----EAREAQTEQAPRE-----TASDQDRSGSGSERDEGED----- 150

QY 122 RLVPRRSAPRKRTARRARAIAAHYVPRGQDGAQAGVDGTVSG 167

DB 151 ---RQSE-----RRSGRR-----RAGDDAQOGQRRSRDG 179

RESULT 10

T497\_MTCUTU

ID T497\_MTCUTU STANDARD; PRT; 310 AA.

AC Q1162;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein RV0497.

GN RV0497 OR MT0517 OR MTCY2069. 23.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

RA

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock R., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;

RA "Elucidating the biology of Mycobacterium tuberculosis from the complete genome sequence."

RT Nature 393:537-544(1998).

RL

RN

[2]

SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RA Fleisemann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Bishal W.;

RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: TO M.LEPRAE M2433.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a statement agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

CC

DR EMBL: Z77162; CAB00923.1; -

DR EMBL: AE006952; AAK44740.1; -

DR TIGR; MT0517; -

DR TubercuList; RV0497; -

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 231 251 POTENTIAL.

FT TRANSMEM 257 277 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT DOMAIN 33 39 POLY-ARG.

FT DOMAIN 197 202 POLY-ALA.

SO SEQUENCE 310 AA; 33092 MW; 4954027F694DF5C2 CnC64;

Query Match

Best Local Similarity 6.4%; Score 93; DB 1; Length 310;

Matches 80; Conservative 25; Mismatches 103; Indels 120; Gaps 18;

QY 14 PLPLRSLSRSDGA---VRQAPPAPMAARRSGRRG-----R 47

DB 6 PETESSGNROISVAVELARQGTGAP--ARRRRRRGSDATVAVELGELPIIRDHHH 63

QY 48 RGEPTALLVPLATGLALACIGLLAVSLGSRASLSAOPAEELVAEE----- 99

DB 64 ACPDAHASQSPANR-----VOYGEAAPSPAPVABQ-VAEPTITVYS 109

QY 100 DQPSLEPNQTESQDAPFLNLRVPRR---SAPRGKTRARRAI--AAHY----- 146

DB 110 QEPWRPKSPQDRRESGELSEYPRPLRHTSDRAPAGPSCALMSDPVEHVDLW 169

QY 147 -----EVPFRGQDS-----AQAGVDGIVSGHEARINSS--PLKYN 182

DB 170 DVLDTVEGEAEAEVREARQAPGGRHAAAGTVDGGAELARVARRALDVPTLMR 229

QY 183 RQGEFIYTR-----AGLYLYCOYHFD-----GRAVYTKLPL-----LVDSGLALR 225

DB 230 ---GALVYLQSTILAVAFGGLF-----IAFDLMRRNSIVLVSVYILGLVSVRAVR 281

QY 226 CLEFSAT-----AASLAPQLRIQ 246

DB 282 KTEIASLTLINAVGALLITLGP-LALLQ 308

RESULT 11

TN14\_HUMAN

ID TN14\_HUMAN STANDARD; PRT; 240 AA.

AC Q43557; Q75476; Q961D2; Q8WVE8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 14 (Herpesvirus entry mediator-ligand) (HVEM-L).

GN TNFSF14 OR LIGHT OR HVEM-L.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RA

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=98122340; PubMed=9462508;

RA Mauri D.N., Ebner R., Montgomery R.I., Kochel K.D., Cheung T.C., Yu G.-L., Ruben S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G., Ware C.F.;



"LIGHT, a new member of the TNF superfamily, and lymphotoxin alpha are ligands for herpesvirus entry mediator.";  
 RT Immunol 8:21-30(1998).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.  
 RC TISSUE-LIVER;  
 RA MEDLINE-98438532; PubMed-9765287;  
 RA Harrop J.A., McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,  
 RA Tan K.B., Dede K., Spampinato J., Silverman C., Hensley P.,  
 RA DiPrinzio R., Emery J.G., Deen K., Elchman C., Chabot-Fletcher M.,  
 RA Truneh A., Young P.R.;  
 RA Herpesvirus entry mediator ligand (HVEM-L), a novel ligand for  
 RT HVEM/TR2, stimulates proliferation of T cells and inhibits HT29 cell  
 RT growth.";  
 RT J. Biol. Chem. 273:27548-27556(1998).  
 [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.  
 RX MEDLINE-21528948; PubMed-11673523;  
 RA Granger S.W., Butrovich K.D., Houshmand P., Edwards W.R., Ware C.F.;  
 RT "Genomic characterization of LIGHT reveals linkage to an immune  
 RT response locus on chromosome 19p13.3 and distinct isoforms generated  
 RT by alternate splicing or proteolysis.";  
 RL J. Immunol. 167:5122-5128(2001).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTR. Binding to the  
 CC decoy receptor TNFRSF6B modulates its effects. Activates NFkB,  
 CC stimulates the proliferation of T cells, and inhibits growth of  
 CC the adenocarcinoma HT-29. Acts as a receptor for Herpes simplex  
 CC virus.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted  
 CC (isoform 1); Cytoplasmic (isoform 2).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/LIGHT  
 CC delta-TM; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO  
 CC FOUND IN THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID  
 CC TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,  
 CC AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR  
 CC NONHEMATOPOIETIC TUMOR LINES.  
 CC -1- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: The soluble form of isoform 1 derives from the membrane  
 CC form by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.4 sequence differs from that shown due to a  
 CC frameshift in position 178.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF036581; AAC39563.1; -;  
 DR EMBL; AF064090; AAC25169.1; -;  
 DR EMBL; AY028261; AAK26160.1; -;  
 DR EMBL; BC018058; AAH18058.1; ALT\_FRAME.  
 DR HSSP; P01375; 4TSV.  
 DR Genew; HGNC:11930; TNFSF14.  
 DR MIM; 604520; -;  
 DR InterPro; IPR003636; TNF-abc.  
 DR InterPro; IPR000478; TNF\_family.  
 DR Pfam; PF00229; TNF; 1.  
 DR ProDom; PD002012; TNF-abc; 1.  
 DR SMART; SM00207; TNF; 1.  
 DR PROSITE; PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE; PS50049; TNF\_2; 1.

KW Cytokine; Cytotoxin; Transmembrane; glycoprotein; signal-anchor;  
 KW Alternative splicing.  
 FT CHAIN 1 240  
 FT CHAIN 783 240  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 58  
 FT DOMAIN 59 240  
 FT SITE 82 83  
 FT DISULFID 154 187  
 FT CARBOHYD 102 102  
 FT VARSPLIC 38 73  
 FT CONFLICT 120 120  
 FT CONFLICT 214 214  
 SQ SEQUENCE 240 AA; 26351 MW; 49DDBF6E1390839 CRC64;  
 Query Match 6.4%; Score 92; DB 1; Length 240;  
 Best Local Similarity 23.7%; Pred. No. 2;  
 Matches 44; Conservative 19; Mismatches 57; Indels 66; Gaps 6;  
 QY 29 QAQPPAPMARRSORRRGRGEGETALVPLAIGLACLGILLAVSLGSRASLSAQ 88  
 DB 16 QTDIPFTRLRSHRQSCSVARVGLGLL-LWGAGLAVGWFLQLHMRLG----- 66  
 QY 89 EPAQEEIYAEEDDPSELNPTQTESQDPAPFLNRLVPRRSAPKGRTRARRAIAHYEV 148  
 DB 67 -----ENW-----TRIPSPAGSMGLDIERSS-----HEV 92  
 QY 149 HPRPGDQAQAGDGTGVSWEARINSSPLRNROI-----GEFVTRAGLY 196  
 DB 93 NPAHNLGANSLSLIG-----SGGPLWETQGLAFLNGLSYHDALVVTAKGY 141  
 QY 197 YLYQGV 202  
 DB 142 YLYSKV 147  
 RESULT 12  
 ID TNF6\_CERTO STANDARD; PRT; 280 AA.  
 AC QGBDNI;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)  
 DE (CD95L protein).  
 GN TNFRSF6 OR FASL OR CD95L.  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 NX NCBI\_TaxID=9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Lymphocytes;  
 RX MEDLINE-21383618; PubMed-11491535;  
 RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RA "Cloning, sequencing, and homology analysis of nonhuman primate  
 RT Fas/FasL ligand and co-stimulatory molecules.";  
 RT Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that  
 CC transduces the apoptotic signal into cells. May be involved in  
 CC cytotoxic T cell mediated apoptosis and in T cell development.  
 CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of  
 CC peripheral tolerance. In the antigen-stimulated suicide of mature  
 CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DcR3  
 CC modulates its effects (by similarity).  
 CC -1- SUBUNIT: Homotrimer (Probable).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
 CC similarity).



CC -1- PFM: The soluble form derives from the membrane form by  
 CC proteolytic processing (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF344847; AKK37606.1; -  
 DR HSSP: P01375; ATSV.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR PRODOM: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS00049; TNF\_2; 1.  
 KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 280  
 FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT CHAIN 129 280  
 FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT DOMAIN 1 80  
 FT MEMBER 6, SOLUBLE FORM (BY SIMILARITY).  
 FT TRANSHEM 81 101  
 FT CYTOPLASMIC (POTENTIAL).  
 FT SIGNAL-ANCHOR (TYPE-IT MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT EXTRACELLULAR (POTENTIAL).  
 FT PRO-RICH.  
 FT POLY-PRO.  
 FT CLEAVAGE (BY SIMILARITY).  
 FT POTENTIAL.  
 FT DISULFID 201 232  
 FT CARBOHYD 183 183  
 FT CARBOHYD 249 249  
 FT CARBOHYD 259 259  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 280 AA; 31407 MW; 729EA60067B7D398 CRC64;

Query Match 6.4%; Score 92; DB 1; Length 280;  
 Best Local Similarity 20.8%; Pred. NO. 2.4;  
 Matches 62; Conservative 44; Mismatches 118; Indels 74; Gaps 14;

OY 13 LPLPSTLSRGGAGVRAQAPAPMARSRQRGRGEGTALVPLA----- 61  
 DB 30 LPPSPSVRRRPGQRPPPPPPPPPP-----PPPPPPPLPLPPPKKGNHST 78  
 OY 62 GLGLALACIGLLAVVSLGRASLSAQEPAGEELVAEEDOPSELPQTEESQDAPPLN 121  
 DB 79 GLCILLVMEFVVLVALVGLG--LGMFQLFHLQEL-----AELRESTSKHTASSLEK 128  
 OY 122 RLVRPRKRAPKGRKTRARARAIATAYEVRPRGQDGAQGVDTGSGMEDA-RINSSPLR 180  
 DB 129 QIGHP-----SPPEKKEQRK--VAHLTGKPNRSMPLE-----WEDYGVIVLSGVK 174  
 OY 181 YNRQIGETIVRAGIYLYCOVHEDEKA-----VYLKLD-----LLVDVILALR 225  
 DB 175 YKK--GGIVINETGLYFYYSKYF--RGOSCTNPLPSHKVYMNSSKYPDOLVMEBKMSY 231  
 OY 226 CLEESATPAASSLGPQLRICOVSGLLALRPSSSLRIRLPMVAHLKAAPFLAYFGLPQV 283  
 DB 232 CTGGMAMHSSVILGAVFNLTSIDHLY-----VNVSELSLVNFEESQ--TFGLYKTL 280

RESULT 13  
 ID\_NFC4\_HUMAN STANDARD; PRT; 902 AA.  
 AC 014934;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nuclear factor of activated T-cells, cytoplasmic 4 (T cell)

DE transcription factor NFAT3) (NF-ATc4) (NF-AT3).  
 GN NFATC4 OR NFAT3.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 CC NCBTaxid:9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-T-cell;  
 RX MEDLINE-9526130; PubMed-7749981;  
 RA Hoey T., Sun Y.-L., Williamson K., Xu X.;  
 RT "Isolation of two new members of the NF-AT gene family and functional  
 RT characterization of the NF-AT proteins.";  
 RL Immunity 2:461-472(1995).  
 RN [2]  
 RP REVIEW.  
 RX MEDLINE-99189746; PubMed-10089876;  
 RA Crabtree G.R.;  
 RT "Genetic signals and specific outcomes: signaling through Ca2+,  
 RT calcineurin, and NF-AT.";  
 RL Cell 96:611-614(1999).  
 CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE  
 CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 AND IL-  
 CC 4 (BY SIMILARITY).  
 CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX  
 CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING  
 CC CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT  
 CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE  
 CC ACTIVATING PROTEIN-1 FAMILY, MAF, GATA4 AND GABP300 CAN ALSO BIND  
 CC TO THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND  
 CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-  
 CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT  
 CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED  
 CC AND TRANSIENT CALCIN SIGNALS. THE SUBCELLULAR LOCALIZATION OF  
 CC NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, LUNG, KIDNEY,  
 CC TESTIS AND OVARY. WEAKLY EXPRESSED IN SPLEEN AND THYMUS. NOT  
 CC EXPRESSED IN PERIPHERAL BLOOD LYMPHOCYTES.  
 CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND  
 CC COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).  
 CC -1- PFM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY  
 CC CALCINEURIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L41066; AAA79175.1; -  
 DR TRANSFAC: T02462; -  
 DR GENEW: HGNC:7778; NFATC4.  
 DR MIM: 602699; -  
 DR InterPro: IPR002809; IPT\_TIG.  
 DR InterPro: IPR000451; NF\_Rel\_dor\_fam.  
 DR Pfam: PF01833; TIG; 1.  
 DR SMART: SM00429; IPT; 1.  
 DR PROSITE: PS01204; REL\_1; FALSE\_NEG.  
 DR PROSITE: PS02544; REL\_2; 1.  
 KM Transcription regulation; Activator; Nuclear protein; DNA-binding;  
 KW Repeat; Phosphorylation.  
 FT DOMAIN 62 69  
 FT DOMAIN 114 119  
 FT DOMAIN 213 293  
 FT REPEAT 213 229  
 FT REPEAT 277 293  
 FT DOMAIN 297 304  
 FT DOMAIN 268 270  
 FT DOMAIN 430 437

POLY-PRO.  
 CALCINEURIN-BINDING.  
 2 APPROXIMATE SP REPEATS.  
 SP 1.  
 SP 2 (APPROXIMATE).  
 POLY-PRO.  
 NUCLEAR LOCALIZATION SIGNAL.  
 DNA-BINDING.

F1 DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.  
 S1 SEQUENCE 902 AA; 95472 MW; E59F15F7647A7C6 CRC64;  
 Query Match 6.4%; Score 92; DB 1; Length 902;  
 Best Local Similarity 25.0%; Pred. No. 9;  
 Matches 48; Conservative 15; Mismatches 73; Indels 56; Gaps 8;  
 Q1 14 PLPRSLGRDGA V RQAPAPMAARRS-----QRRRGRRGEPGALLVPLALGLALAC 69  
 D1 236 PPSGGRPEDESWLLLSAPGTPASPRSPGCKRRYSSSGTPSSA-----SPALSR 286  
 Q2 70 LGILLAVSISGRASLSAQPEAEELVAEDQDPELNPQTESQDAPFLNRLVRRRS 129  
 D2 287 RG-----SLGEEGS-----EPPPPPL-PLARDGSGSPEDYVGAAPAES 325  
 Q3 130 AP-KGRTRARRAIA-----AHYEVRPRGQCAQAGVD-----GVVSG 167  
 D3 326 IPQTRRTSSQAVALLPRSEBPASCKGLPLGAEESVAPPGSKREYAGMDYLAVPSPLA 385  
 Q4 168 WEERINSSSPL 179  
 D4 386 WSKRIGHSPI 397  
 RESULT 14  
 MLH1\_MOUSE STANDARD; PRT; 760 AA.  
 AC Q9JK91; 062454;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA mismatch repair protein Mlh1 (Mult. protein homolog 1).  
 GN MLH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 FN [1]  
 FP SEQUENCE FROM N.A.  
 FC TISSUE=Testis;  
 RA Kumaran M., Rao M.R.S.;  
 RT "Cloning of the cDNA of the Mult. homolog, MLH1 from mouse testis.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE OF 1-151 FROM N.A.  
 RX MEDLINE=96270514; PubMed=8674118;  
 RA Edelmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,  
 RA Umar A., Kunkel T., Caltonetti G., Chaganti R., Pollard J.W.,  
 RA Kolodner R.D., Kucherlapati R.;  
 RT "Meiotic pachytene arrest in MLH1-deficient mice.";  
 RL Cell 85:1125-1134(1996).  
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.  
 CC -1- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MTHL/HEB FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF250844; AAF64514.1; -;  
 DR EMBL: U60872; AAC52672.1; -;  
 DR EMBL: U59881; AAC52672.1; JOINED.  
 DR EMBL: U59882; AAC52672.1; JOINED.  
 DR EMBL: U59883; AAC52672.1; JOINED.  
 DR EMBL: U59884; AAC52672.1; JOINED.  
 DR HSSP: P23367; 1BK1.  
 DR MGI: 101938; Mlh1.  
 DR InterPro: IPR003594; ATPbind\_ATPase.

DR InterPro: IPR002099; DNA\_mis-repair.  
 DR Pfam: PF01119; DNA\_mis-repair; 1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR TIGRfam: TIGR00585; mult; 1.  
 DR PROSITE: PS00058; DNA\_MISMATCH\_REPAIR\_1; 1.  
 DR KWA repair.  
 S1 SEQUENCE 760 AA; 84679 MW; 173C809372A29186 CRC64;  
 Query Match 6.3%; Score 90.5; DB 1; Length 760;  
 Best Local Similarity 22.7%; Pred. No. 9.8; 126; Indels 81; Gaps 14;  
 Matches 70; Conservative 31; Mismatches 126; Indels 81; Gaps 14;  
 Q1 13 PLPRSLGRDGA V RQAPAPMAARRS-----QRRRGRRGEPGALLVPLALGLG 64  
 D1 436 LPAPAEAAASENLERISMETSDAQKAPATSPSSRRNRH--EDSDVEAVENMSGKE 493  
 Q2 65 LALAC-----LGILLAVSISGRASLSAQPEAEELVAEDQDPELNPQ--TEESODPA 117  
 D2 494 MTACACYPRRRIINTSLSLQEEISERCHETLRE--ILRNHSFVGCVNPQMALAHOQTKL 551  
 Q3 118 PFLNRLVRRRSPAPKGRTRARRAIAHAYEV-----HPRPGDGAQAGVDG 163  
 D3 552 YLLN-----TTKLSSELFYQILLYDPANFGVLLSEPAFLDLAMALDS 596  
 Q4 164 TVSGWEERINSSSPLRYNRQIGFIVTRAGLYLYCQVHFDEGKAYTLKLLVDG--- 220  
 D4 597 PESGWTDDGPKGLAEY---IVEFLKKAEMLADFSVEIDE-EGMLIGLPLIDSYVP 652  
 Q5 221 -----VLAR-----CLEFESATAAS--SLGP--LRICOVSGLLALRPGSS 258  
 D5 653 PLGGLPFLIRLATEVNMDEKCFESLSKCAFYSIRQYILLESTLSGQSDMPGST 712  
 Q6 259 LRIRTLPM 266  
 D6 713 SK-----PW 716  
 RESULT 15  
 CADF\_HUMAN STANDARD; PRT; 814 AA.  
 ID CADF\_HUMAN  
 AC P55281;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Muscle-cadherin precursor (M-cadherin) (Cadherin-15) (Cadherin-14).  
 GN CDH15 OR CDH14 OR CDH3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 FN [1]  
 FP SEQUENCE FROM N.A.  
 RC TISSUE=skeletal muscle;  
 RX MEDLINE=97184182; PubMed=9030594;  
 RA Shibata T., Shimoyama Y., Gotoh M., Hirohashi S.;  
 RT "Identification of human cadherin-14, a novel neurally specific type  
 RT II cadherin, by protein interaction cloning.";  
 RL J. Biol. Chem. 272:5236-5240(1997).  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC -1- FUNCTION: CADHERINS INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. M-CADHERIN IS PART OF THE  
 CC MYOCENT PROGRAM AND MAY PROVIDE A TRIGGER FOR TERMINAL MUSCLE  
 CC DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:13:22 ; Search time 15.6228 seconds  
(without alignments)  
1747.586 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLLPRPSIG.....PWAIKKAFLTYFGFQVH 284

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106.5	7.4	254	2	4-1BB ligand - hum
2	104	7.2	1323	2	N-methyl-D-aspartate
3	100.5	7.0	375	2	hypothetical prote
4	99.5	6.9	378	2	H87333
5	99	6.9	878	2	HLyD family secret
6	98	6.8	206	2	T17245
7	97	6.7	441	2	T34961
8	93.5	6.5	776	2	S41710
9	93.5	6.5	1657	2	T36946
10	93	6.4	310	2	T15838
11	92	6.4	351	2	D70745
12	92	6.4	566	2	B34768
13	91.5	6.3	212	2	T35203
14	91.5	6.3	664	2	A70611
15	91	6.3	681	2	D83231
16	91	6.3	1217	2	H83044
17	90.5	6.3	814	2	T00270
18	90.5	6.3	1329	2	G02878
19	90	6.2	660	2	D87226
20	89.5	6.2	886	2	T03038
21	88.5	6.1	210	2	S07132
22	88.5	6.1	278	2	D87394
23	88.5	6.1	281	2	A49266
24	88.5	6.1	998	2	T38707
25	88	6.1	531	2	G83022
26	87.5	6.1	492	2	C83153
27	87	6.0	197	1	A87471
28	87	6.0	204	1	JH0309
29	87	6.0	445	2	S24641
					T30604

30	87	6.0	2124	2	H83357	probable non-ribos
31	86.5	6.0	399	2	F83633	hypothetical prote
32	86.5	6.0	439	2	B70629	hypothetical prote
33	86.5	6.0	755	2	B75346	probable competent
34	86.5	6.0	762	2	E98121	hypothetical prote
35	86.5	6.0	764	1	S14113	hypothetical prote
36	86.5	6.0	810	2	D95256	1-phosphatidylinos
37	86.5	6.0	887	2	AC0535	Arp-dependent Clp
38	86	6.0	139	2	B69953	hypothetical prote
39	86	6.0	565	2	G98331	probable oligopept
40	86	6.0	568	2	E83325	probable chemotaxi
41	86	6.0	777	2	A87309	hypothetical prote
42	86	6.0	839	2	F75518	hypothetical prote
43	86	6.0	933	1	B48349	glycoprotein B pre
44	85.5	5.9	563	2	C98227	hypothetical prote
45	85.5	5.9	563	2	A83059	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

138427

4-1BB ligand - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000

C:Accession: I38427

R:Allderson, M.R.; Smith, C.A.; Tough, T.W.; Davis-Smith, T.; Armitage, R.J.; Falk, B.

Eur. J. Immunol. 24, 2219-2227, 1994

A:Title: Molecular and biological characterization of human 4-1BB and its ligand.

A:Reference number: I38426; MUID:94374434; PMID:8088337

A:Accession: I38427

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-254 <RES>

A:Cross-references: EMBL:003398; MID:9571322; PIDN:AA5134.1; PID:9571323

Query Match 7.4%; Score 106.5; DB 2; Length 254;

Best Local Similarity 26.5%; Pred. No. 0.32;

Matches 74; Conservative 31; Mismatches 93; Indels 81; Gaps 13;

QY	32	PPAPMAARRSQRRRGRGEGTALVPLALGIALACGLIAVSL-GSRASL-SAOE	89
DB	16	PPAP-----PAPACRVLP-MALVAGLLLLLAACAVFLACPMVAAGASAPSSAAS	67
QY	90	PAQELVAREDDPSSELNQTESQDPAFLNIVAPRRSAPKGRTRARBAIAHYEVH	149
DB	68	PLURE-----GPELSP-----DDPAGLIDL-----RGMFAQLVAQNVTL-	102
QY	150	PRPGDQAQGVYGVSGWEE--ARINSSPLRYNRQIGEFIVTRAGLYIYICQVHDE	206
DB	103	-----IDGPLSWSDPGLAGVSLGGLSTYEDRKELVNAAGYIYFFQ-----	146
QY	207	GAAYVTKLDLVDG-----VIALRCLEPSSATASSLGPQLRCOVSG-----	249
DB	147	-----LELRVYVAGSGSGVSLALHQLPLRSAGAAALLVTDLPASSEARNSAFQOG	201
QY	250	LLALRPSSLRIRTLPMWAIKKAAPFL-----TYFGFQVH 283	
DB	202	RLHLSAGQRLGVHTEARAHAWOLTCGATVIGLFRV 240	

##### RESULT 2

S27224

N-methyl-D-aspartate receptor epsilon-4 chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 25-Feb-1994 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999

C:Accession: S27224

R:Iikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.; Watanabe, M.; Inoue, Y.

FEBS Lett. 313, 34-38, 1992

A:Title: Cloning and expression of the epsilon-4 subunit of the NMDA receptor channel

A:Reference number: S27224; MUID:93050214; PMID:1385220

A:Accession: S27224

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1323 <IEX>  
A:Cross-references: EMBL,D12822  
C:Superfamily: N-methyl-D-aspartate receptor 2D; glutamate receptor homology  
C:Keywords: transmembrane protein  
F:451-879/Domain: glutamate receptor homology <GnH>

Query Match	7.2%	Score 104;	DB 2;	Length 1323;
Best Local Similarity	23.7%	Pred. No. 3.2;		
Matches	56;	Conservative	21;	Mismatches 97;
				Indels 62;
				Gaps 8

```

QY 10 ARRLPLRSLGSRGCGVVRQAQPPAPM-----AAKRSRRRGRCGPECTALL----- 57
Db 903 AKPPPPQQLPSPLVYPAAPPPGCGAPFVPPERRAAADRMWRRAAG-TGPPGAGALADGFHRY 961
QY 58 -----PLALGIGLALAC-LGILLAVVSLGSRASLSAQEPAQEBLVAEEQDQDSE----- 105
Db 962 YGPIEPQGLGGEARAPRPG--AAGRPPLSPITTPQPPQKPPSYFLVAYEQEPAEPPAGAF 1019
QY 106 -----LNPT-----EESODPAPELNLVPRRS-----A 130
Db 1020 GGPSPAPAPAAAAAAGPPLCLRLAEDEDSPPASAGRLVTPRRAARCWCGAARAALGPR 1079
QY 131 PKGKTRARRAIAAHAYEHVHPFGQDGAQAQGVQGVTSQGWEEARINSSSPLRTNRQIG 186
Db 1080 PHHRRVATAPPCAYLDLESPESDSDSGLGASTGLGLEPWWFADFPPYPAERIG 1135

```

RESULT 3  
A75264  
hypothetical protein - *Deinococcus radiodurans* (strain RJ)  
::Species: *Deinococcus radiodurans*

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: A75264  
R:Wolfe, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Mc  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
A:Reference number: A75250; MIMD:20036896; PMID:10567266  
A:Accession: A75264

Query Match	7.0%;	Score 100.5;	DB 2;	Length 375;
Best Local Similarity	24.0%;	Pred. No. 1.5;		
Matches 69;	Conservative 27;	Mismatches 84;	Indels 107;	Gaps 13

QY 25 GAVRQAOPAPMA-----RRSORRRGRGEPGT--ALLVPLALGLIALACTILLAVV 77  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 5 GRGHRCPTGRATARLRLRRPGRAPCRGCGHPAARGRLRLDIALALA-----LYA 58  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 78 SLGSRASLSAQE-PAQBELYAEEODPSSELPQTSESDAPFTLNRLVPRRSAPKGRKT 136  
:  
Db 59 DVTTRPFLPAGCLPTLLIDLAR-----ODLESLIDSP----- 89  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 137 RARRATAAHVEHVR--PGGDGAQAG--VDGTSGMEGARINSSPLRYNQIG-----F 188  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 90 TDRLAQARAAATPRLPBGDDGAGFGHPISITLAGE--RLSAASLRAMHARFGORGVILG 147  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 189 IYTRAGLIYLIVY-----COVHEDEGRAVYLKLIDLVGYALRCLEEFS 231  
: : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 148 IVAADAAOEYELVLACGLFADWDGPEDRPMFAHFOPGRHL----- 168  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 232 ATTAASSLGPOLRLCOVS-----GLIAIRRESSLRIRK 263  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Db 189 ---PSATGEQTHLSLVAPGPGPRDPDWLPWQLALTALSGCSASRLFT 232

RESULT 4  
H87333

HlyD family secretion protein [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: H87333  
 R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: H87333  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-378 <SMO>  
 A:Cross-references: GB:AE005673; MID:913421902; PIDN:AAK22668.1; GSPDB:GN00148  
 C:Genetics:  
 ;:Gene: CCO683

Query Match	6.9%	Score 99.5;	DB 2;	Length 378;
Best Local Similarity	25.1%;	Pred. No. 1.8;		
Matches 73;	Conservative 29;	Mismatches 106;	Indels 83;	Gaps 13

```

QY 19 LSRDQGAVRQAOPAPMAARRSQRRGR--GEBGTALIVPALIGLALACTIGLLA 75
Db 36 LIAARGGGIYANVH-----REGDVRVVVGGP IATLTLSLLEGDSFAVLISR-- 81
QY 76 VVLSGRASLSAOEPA--QELVAEEDODPSSELNPTQESOPAPFLMRIRPRPS--AP 133
    |||:::|
Db 82 --SLGASGSAAGSRAATQALAAESRQ-----LTORRALNRELAKRSIALO 129
QY 132 KKKRTRAR-----RAIAHYEVHPRPGDDQAQVGDVTSQWMBEARIINSSPLRYNRQI 185
    |||:::|
Db 130 TERTLLAARAEVARAETTLAAGFLSPRE--LQARRSARLVAQDEASTLSQALSYEROI 185
QY 186 GEFTYTRAGLYLYLCOVHFEDEGRKAVYLKIDL---LVDGVLAVRCIEEFSAPAAASLSLPOL 242
    ||:::|
Db 186 GEV-----DARLAIAPIDLQAARAEAASTLAGLEQ-QAIQVQENGRIV 227
QY 243 RLCOVSSGLA-----LRPSSLRIRLTPAHAKKAAPFL 275
    |||:::|
Db 228 VYATYGVAAALPVEACQTYGVGAANAVALTPDGSALVAMEL-YAPRRAAGFV 277

```

RESULT 5  
T17245  
hypothetical protein DKFZp586J0917.1 - human (fragment)

C:Species: homo sapiens (man)  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: 117245  
 A:Accession: 117245  
 R:Koelner, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, September 1999  
 A:Reference number: Z18722  
 A:Accession: 117245  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-878 <KOE>  
 A:Cross-references: EMBL:AL117455  
 A:Experimental source: adult uterus; clone DKFZp586J0917  
 C:Genetics:  
 D:Note: DKFZp586J0917.1

```

Query Match      6.9%  Score 99  DB 2  Length 878;
Best Local Similarity 24.2%  Pval. NO. 5.1,
Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6

11 RRLPLPRSLGSRDGGAVR-----QAQPRAPMAR-----RSQRRGR- 47
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
327 KQISPAEDLEFDGGGCGQGVNDGLNRELGNQREAPRGLQONPOVLLMDEQRLAGRG 386

```

Dy 11 RRLPLRSLGSDGGAVR-----QAQRPAPMAR-----RSQRRRGR- 47  
||:|||:|:  
Db 327 RQLPSAEDLEFDGGGPVDDGLHRELGHQPPEAGSPARLPLOHPVLMEQQRLAGRL 386



submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of *C. elegans* cosmid C54D2.  
 A:Reference number: Z18415

A:Accession: T15838

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1657 <MIN>

A:Cross-references: EMBL:U03548; NID:g1017804; PID:g1017809; PIDN:AAV9201.1; CESP:C54D2

A:Gene: CESP:C54D2.5

A:Insertions: 40/2; 67/3; 86/3; 121/2; 199/3; 230/2; 308/2; 334/2; 370/2; 439/3; 470/3; 513

C:Superfamily: sodium channel protein

Query Match 6.5%; Score 93.5; DB 2; Length 1657;  
 Best Local Similarity 23.2%; Pred. No. 29;  
 Matches 63; Conservative 38; Mismatches 99; Indels 71; Gaps 12;

OY 52 GTALIVPLAIGLALACLG-----LLAVVSLGSRAS-----LSAOEPAOEELVAED 100  
 DB 881 GMAQTNPMAALYFVALTFGVYLFNLVAILVGEFQESKEERKOLEEDARKQAVEED 940  
 OY 101 QDPSELNPOTESODPAPFLNRLVLP-----RRSAPKGRTRARRAIAHYEVRPPQ 154  
 DB 941 ERKELELIIAKTSPA--FNNGVAPAECTCQRPSP--ESPRLISANT--HSPER 994  
 OY 155 DGAQAGVDGTVSGWEARINSSPL-----RYNRQIGEFIVTRAGLYYLCOV 202  
 DB 995 KHS-ANLDAITD--KRLVLRNSAPFDRSPVSEGRDRLNRHNASLVPVANGVYRRQVR 1051  
 OY 203 H-----FDEKAYLIKLDLYDGLALRCLE---EFSATPAASSLIGPOLRLCOV 247  
 DB 1052 HSMKASQELKQALAEERKREAKONTVRLKTKLHNRTFS----- 1095  
 OY 248 SGLLALRPGSSLRITLPMALHKAAPLTYF 278  
 DB 1096 --LPLMGPKNPLRIKCLQTQKWFDTYVLF 1124

#### RESULT 10

hypothetical protein RV0497 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: D70745

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; M0ID:98295987; PMID:9634230

A:Accession: D70745

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-310 <COL>

A:Cross-references: GB:277162; GB:AL123456; NID:g3261606; PIDN:CAB00923.1; PID:g255036;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0497

Query Match 6.4%; Score 93; DB 2; Length 310;  
 Best Local Similarity 24.4%; Pred. No. 4.8;  
 Matches 80; Conservative 25; Mismatches 103; Indels 120; Gaps 18;

OY 14 PLPRSLGSRDGA---VROAOPAPMAARRSORRG-----R 47  
 DB 6 PERESSNRKQISVALLARQGVGAP--ARRRRRRRDSALTYAELTGEPITRDDHH 63  
 OY 48 RGEPTALVPLAIGLALACIGLLAVVSLGSRASLSAOEPAOEELVAED----- 99  
 DB 64 AGPDNAHSQSPANR-----VOGEAAPQSAPEPAED--VADEPRTTYMS 109  
 OY 100 DQDSELNPOTESODPAPFLNRLVRRP-----SAPKGRKTRARRAI--AAHY----- 146

DB 110 QPERBMPSPPODRRESGPSELSEYRPLRTHSDRAPAGPPSGAEHMSPPVEHYDPLMW 169  
 OY 147 -----EVHPRPGDG-----AQAGVCTVSGWEARINSS-----PLRYN 182  
 DB 170 DVIDTEVGEAEATEVEADPGRGERHAAAAAGTDEGGGAARVARRALDVPPLMR 229  
 OY 183 ROIGEFIVTR-----AGLYLYCOVHPE-----GKAVYKLDL-----LYDGVLAIR 225  
 DB 230 ---GALVLOSILAVAGAGLIF-----IAPDLMRMSIYALVSVVILGLVYSRAVR 281  
 OY 226 CLEEFSAF-----AASSLGPOLRLCO 246  
 DB 282 KTEIDIASTLIAVANGALITLGP--LALLQ 308

#### RESULT 11

B34768

ORF5 protein - Orf virus (strain NZ2)

C:Species: Orf virus

C>Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 08-Oct-1999

C:Accession: B34768

R:Fraser, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.

Virology 176, 379-389, 1990

A:Title: Sequence analysis of the inverted terminal repetition in the genome of the p

A:Reference number: A34768; M0ID:90266454; PMID:2129563

A:Accession: B34768

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <FRA>

A:Cross-references: GB:M30023; EMBL:M37623; NID:g332561; PIDN:AAA46788.1; PID:g332564

Query Match 6.4%; Score 92; DB 2; Length 351;  
 Best Local Similarity 23.0%; Pred. No. 6.6;  
 Matches 45; Conservative 18; Mismatches 73; Indels 60; Gaps 7;

OY 14 PLPRSLGSRDGAVRQAOPAPMAARRSQ--RRGRGREGPTALLVPLAIGLALACGL 72  
 DB 211 PLPRRAR--GRORRGOPPPRRARRAQOPRRAPRAAG----- 247  
 OY 73 LIAVSLGSRASLSAOEPAOEELVAEDDOPSELNPOTESODPAPFLNRLVRRSAPK 132  
 DB 248 -----ARRGRGAPRQOQOPRVORAAAQRRARQOROP 283  
 OY 133 GKTRARRAIAHYE-VHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVT 191  
 DB 284 PRVRRARRARRAQQRHQR--RRGRARRRCS-----RVYSKD-----SREVGWCCD 332  
 OY 192 RAGLYLYCOVHPEDEG 207  
 DB 333 KERRYIRVLLHPEDEG 348

#### RESULT 12

T35203

probable two-component sensor - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 31-Jan-2000

C:Accession: T35203

R:Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z21571

A:Accession: T35203

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-566 <SEE>

A:Cross-references: EMBL:AL022374; PIDN:CAA18527.1; GSPDB:GN00070; SCOEDB:SC5B8.19C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC5B8.19C

C:Superfamily: two-component sensor histidine kinase; sensor histidine kinase homolog

Query Match 6.4%; Score 92; DB 2; Length 566;





```

QY 105 ELNPTSESDPAPFLNRLVNPRRSAKGRKTRARRAAIAHVEYHNPREDGQAQAGYDGT 164
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
pB 549 --NSOPGGLRPAP-----TPPAQV-----WLOKRPGRPAQOL---GR 583
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 165 VSGNEEARINSSPLRYN--RQIEFLVTRAGLYLYCYQHPEGRKVVYLKLD---LL 217
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
pB 584 TSGWVH-----RAHLRHHVAMMG-----GYEYL-----KIDE-KGLIRIVGVERWLE 626
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 218 VDGVL-----ALRCLEEFSAATAASLGPOLRL-----COVSGSL-----ALRDGSSLRIR 262
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
pB 627 VDNVVICAGQEPLRELDQIROTES-----LRFPLIGCANVAGELDAKRAIRFGAMLAAR 660
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

Search completed: May 8, 2003, 02:22:06  
Job time : 18.6228 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - nucleic search, using frame\_plus\_p2n model

```
Run on:      May  8, 2003, 02:22:47 ; Search time 210.35 Seconds
```

3040.496 Million cell updates/sec

Title: US-09-245-198A-4

Sequence: 1 MSLLD~~FEISAR~~RLPLPSLG.....PWAHLKAAPFLTYEGLFQVH 284

**Scoring table:**

Xgapop	10.0	,	Xgapext	0.5
Ygapop	10.0	,	Ygapext	0.5
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Maximum DB seq Length: 20000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Command line parameters:

```

MODEL=irmme_p2n.model -DEV-xip
-O/-crg12/USPTO.spool/US09245159/runtat.06052003.170239.9605/app.query.fasta_1.846
-db-n Genesecv.101002 -OPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIST=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09245159.ecgc1.1.323.erna1.06052003.170239.9605 -NCPU=6 -CPU=3
-NO_XLEXY -NO_MAMP -LAREJOURNEY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -IGAPOP=10 -IGAPEXT=0.5 -DELLO=6 -DELEXT=7

```

Database : N\_Geneseq\_101002:\*

1:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1980.DAT *
2:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1981.DAT *
3:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1982.DAT *
4:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1983.DAT *
5:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1984.DAT *
6:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1985.DAT *
7:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1986.DAT *
8:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1987.DAT *
9:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1988.DAT *
10:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1989.DAT *
11:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1990.DAT *
12:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1991.DAT *
13:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1992.DAT *
14:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1993.DAT *
15:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1994.DAT *
16:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1995.DAT *
17:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1996.DAT *
18:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1997.DAT *
19:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1998.DAT *
20:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA1999.DAT *
21:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA2000.DAT *
22:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA2001A.DAT *
23:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA2002.DAT *
24:	/SID52/gcgdata/geneseq/geneseqn -emb1/NA2002.DAT *

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1444	100.0	1373	19	AAV18600	Homo sapiens tumour
2	1406	97.4	1364	24	ABK34881	Human cDNA encoding
3	1377	95.4	1421	20	AAV56000	Human tumour necro
4	1366	94.6	1353	21	AAA49717	Human PMO207 cDNA
5	1366	94.6	1353	24	ABK40255	CDNA encoding huma
6	1265	87.6	1236	19	AAV47613	TNF related endothe
7	1265	87.6	1236	22	AAV04350	Human TREPA (TNF r
8	1066	73.8	898	22	AAV03964	Expression vector
9	1062	73.5	1030	20	AAV23424	Human TRRL3 DNA.
10	1020	70.6	1168	19	AAV18599	Mus musculus tumou
11	968	67.0	701	20	AAV23425	Mus musculus tumou
12	127	8.8	2856	19	AAV21440	Homo sapiens BAROL
13	126.5	8.8	1560	21	AAA49205	Corn putative leci
14	124.5	8.6	2832	20	AAV65289	Murine JNK-interact
15	124.5	8.6	154746	24	AAV25519	Human herpesvirus
16	124	8.6	3957	22	AAA09686	HSV-2 immediate ea
17	124	8.6	154746	24	AAV25519	Human herpesvirus
18	123.5	8.6	11705	19	AAV62160	HSV-2 strain SB5 C
19	123.5	8.6	117213	19	AAV62160	HSV-2 strain SB5 C
20	120.5	8.3	2828	21	AAV64674	CDNA of a human di
21	120.5	8.3	2828	24	ABK64503	Human benign prost
22	120.5	8.3	2828	24	ABL61804	Colon adenocarcinoma
23	117.5	8.1	3860	24	ABL67211	Thyroid cancer rel
24	117.5	8.1	3501	17	AAV42224	Human TATRA-binding
25	117.5	8.1	3501	18	AAV79598	TATRA-binding prote
26	117.5	8.1	3502	15	AAV07338	DNA encoding an ac
27	117	8.1	1030	20	AAV99230	Nucleotide sequenc
28	117	8.1	1030	21	AAV72514	Drosophila melanog
29	116	8.0	978	23	ABL21473	Chicken encoding tum
30	116	8.0	2148	24	ABK11680	Chicken beta-actin
31	115	8.0	1275	11	AAQ01620	Expression vector
32	115	8.0	1344	12	AAQ13318	Hybrid promoter o
33	115	8.0	1345	11	AAQ03062	Chicken beta-actin
34	115	8.0	1495	22	AAE84455	Mouse Meg1/Crbl1 n
35	115	8.0	4211	22	AAH21793	Rat expression prot
36	115	8.0	5504	24	AAAL43172	Human expression p
37	115	8.0	5565	24	AAAL43171	Hprt gene containi
38	115	8.0	5581	22	AAH41035	Positive selection
39	115	8.0	5759	24	ABK49521	Positive selection
40	115	8.0	6148	24	ABK49520	PCAGS-1/SEPOR2 rat
41	115	8.0	6256	24	AAAL43173	Human PCSK1-alpha
42	115	8.0	8033	22	AAV12739	Human strain SB5 C
43	114.5	7.9	12700	19	AAV61133	Vitamin D hydroxyl
44	113.5	7.9	3836	15	AAO55997	Platenolide syntha
45	113.5	7.9	44377	18	AAV78508	

```

FT      /*tag- a
XX      /note- "tumour necrosis factor related ligand"
XX      MO9805783-A1.
XX
XX      12-FEB-1998.
XX
XX      07-AUG-1997: 97MO-US13945.
XX
XX      18-MAR-1997: 97US-0040820.
XX      07-AUG-1996: 96US-0023541.
XX      18-OCT-1996: 96US-0028515.
XX
XX      (BIOJ ) BIOGEN INC.
XX      (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
XX      Browning JL, Chicheportliche Y;
XX
XX      WPI, 1998-145619/13.
XX      P-PSDB; AAW47525.
XX
XX      Tumour necrosis factor related ligand - useful for, e.g. treating
XX      cancer, auto-immune disease and immune responses to tissue grafts
XX
XX      Claim 2; Pages 48-50; 69pp; English.
XX
XX      The sequence is that encoding human tumour necrosis factor related
XX      ligand (TRELL). TRELL or active fragments can be included with a
XX      carrier in pharmaceutical compositions to treat cancer, auto-immune
XX      diseases or immune responses to tissue grafts, or to stimulate or
XX      suppress the immune system. It is useful to screen for TRELL
XX      receptors, by labelling with a detectable label and screening
XX      compositions for binding. Agents interfering with TRELL-receptor
XX      binding can also be screened for, can then be administered,
XX      optionally with interferon- gamma, to induce cell death or
XX      treat, suppress or alter immune responses (especially involving human
XX      adenocarcinoma cells) involving a signal pathway between TRELL and its
XX      receptor. The DNA sequence can be used in gene therapy for
XX      TRELL-related disorders in mammals (especially humans), e.g. tumours,
XX      auto-immune and inflammatory diseases or inherited genetic disorders,
XX      by introducing into cells, and expressing, therapeutically effective
XX      amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
XX      It may also be of use in the preparation of prepare probes for
XX      screening natural/synthetic DNAs for TRELL-encoding sequences
XX      and for antisense therapy.
XX
XX      SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 9.06e-96 Length: 1373
XX      Score: 1444.00 Matches: 284
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 19 Gaps: 0
XX
XX      US-09-245-198a-4 (1-284) x AAV18600 (1-1373)
XX
XX      1 MetSerLeuLeuAspPheGluIleSerAlaArgArgLeuProLeuProArgSerLeuGly 20
XX      1 ATGTCATTTGTTAGACTTTGAAATTTCCGCCGCCGCTCCCTCCCGATCCCTCGGG 60
XX
XX      21 SerArgAspGlyGlyAlaValArgGlnAlaGlnProProAlaProMetAlaAlaArgArg 40
XX      1 TCCCGGATGGGGGGGGGGGAGGACAGGACAGCCGCCGCCCAATGAGCGCGCTCGG 120
XX
XX      41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAla 60
XX      1 AGCCAGAGGCGGAGGCGGCGCGGAGGCGGAGCGGAGCGGCGGCGGCGGCGGCGG 180
XX
XX      61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValAlaValSerLeuGly 80
XX      1 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 240

```

```

XX      81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuAlaGluGluAsp 100
XX      1 AGCCGGGCGATGCTGTCTCCGCCAGGACCTCCAGGAGAGAGCTGTGTGGAGAGAGAG 300
XX
XX      101 GlnAspProSerGluLeuAsnProGlnThrGluLeuSerGlnAspProAlaProPheLeu 120
XX      1 CAGGACCCGTCGGAAGTGAATCCAGACAGAAAGAACCCAGATCTCTGGCGCTTCTCG 360
XX
XX      121 AsnArgLeuValArgProArgArgSerAlaProGlyGlyArgGlyThrArgAlaArgArg 140
XX      1 AACGACTAGTTGGCTCGCAGAGTGCACCTTAAAGCGGAGAAACAGGGCTCGAGA 420
XX
XX      141 AlaIleAlaAlaHisTyrGluValAlaHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
XX      1 GCGATCCAGCCCATTTATGAATTCATCCAGACCTGAGACGAGAGCGAGCGAGCT 480
XX
XX      161 ValAspGlyThrValSerGlyTyrPgluGluAlaArgIleAsnSerSerProLeuArg 180
XX      1 GTGAGCGGAGACAGTGAAGTGGCTGGAGGAGGAGCCAGAAATCAACAGCTCCAGCCCTCG 540
XX
XX      181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
XX      1 TCAACCGCCAGATCGGGAGATTATAGTCACCGGGCTGGCTCTACTACTACTACTGT 600
XX
XX      201 GlnValHisPheAspGlyGlyAlaValAlaTyrLeuLeuLeuAspLeuValAspGly 220
XX      1 CAGGTGCACTTTGATGAGGAGGAGGAGGCTGTACTGAAAGCTGAGCTGTGTGTGTGT 660
XX
XX      221 ValLeuAlaLeuAlaGlyGlyLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyPro 240
XX      1 GTGTGGCCCTCGCTCGCTCGGAGGAATTCACCCAGCTGGGCGCAATCCCTCGGGGCC 720
XX
XX      241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260
XX      1 CAGTCCGCGCTCTCCAGAGTGTCTGGGCTGTGGCCCTCGGCGGCGGAGGCTCTCTCG 780
XX
XX      261 IleArgThrLeuProTyrPheAlaHisLeuAlaAlaProPheLeuThrTyrPheGlyLeu 280
XX      1 ATCCGACACCTCCCTGGGCGGCGGAGGCTCAAGGCTCCCTCCCTCCCTCCCTCCGACTC 840
XX
XX      281 PheGlnValHis 284
XX      1 TTTCCAGGTTCCAG 852
XX
XX      DB: 841 TTCAGGTTCCAG 852
XX
XX      RESULT 2
XX      ID ABR34881 standard; cDNA; 1364 BP.
XX      AC ABR34881;
XX      DT 08-MAY-2002 (first entry)
XX      DE Human cDNA encoding secreted protein #19.
XX      KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
XX      viral infection; bacterial infection; fungal infection; diabetes; asthma;
XX      autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
XX      autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
XX      Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
XX      coagulation disorder; inflammatory disorder; Crohn's disease; incision;
XX      tissue regeneration; wound healing; burn; haematopoiesis;
XX      myeloid cell deficiency; lymphoid cell deficiency.
XX      OS Homo sapiens.
XX      PN WO200177288-A2.
XX      PD 18-OCT-2001.
XX
XX      29-MAR-2001; 2001WO-US10224.
XX      PF 06-APR-2000; 2000US-195582P.
XX
XX      PR

```

XX (GEMV ) GENETICS INST. INC.  
PA

PI Wong GC, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;  
PI Gulukota K, Graham JR;

DR WPI: 2002-179321/23.

PT Five hundred and ninety two polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders

XX Claim 1: Page 82: 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins. The polynucleotides can be used as probes for the  
CC identification and isolation of full length cDNA and genomic DNA. The  
CC polynucleotides and proteins can also be used as nutritional supplements.  
CC The proteins are useful in the treatment of various immune deficiencies  
CC and disorders such as viral infections, bacterial infections, fungal  
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
CC and conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment of  
CC burns, incisions and ulcers. The proteins are also useful for regulating  
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention.

XX Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other:

Alignment Scores:

Pred. No.: 5,08e-93 Length: 1364  
Score: 1406.00 Matches: 276  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.37% Indels: 0  
DB: Gaps: 0

US-09-245-198a-4 (1-284) x ABK34881 (1-1364)

OY 9 SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg 28  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 5 TCCGCGCGGCGCTCCCTCCCGATCCTCGGGATCGGGGGGCGGTGAGG 64  
OY 29 GlnAlaGlnProProAlaProMetAlaAlaArgArgSerGlnArgArgGlyArgArg 48  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 65 CAGGCACAGACCCCGCCCGCCCATGGCGCGCTCGAGCCAGAGGGGGGCGCGG 124  
OY 49 GlyGlnProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAla 68  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 125 GGGGAGACCGGGGACCGCTGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGG 184  
OY 69 CysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGln 88  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 185 TCCCTCGGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244  
OY 89 GluProAlaGlnGlnGlnLeuValAlaGlnGlnAlaGlnAlaGlnAlaGlnAlaGln 108  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 245 GAGCCTGCCAGAGAGAGTGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304  
OY 109 GlnThrGlnGlnSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArgArg 128  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 305 CAGACAGAAAGAAAGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 364  
OY 129 SerAlaProGlyArgGlyThrArgAlaArgArgAlaAlaAlaAlaAlaAlaAlaAla 148  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 365 AGTGCACCTAAAGCGCGGAAACAGCGGGCTGCAAGAGCGATGCGAGCCCATATGAAATT 424

OY 149 HisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyr 168  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 425 CATCCAGACAGCTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484  
OY 169 GluGlnAlaArgGlnLeuAsnSerSerProLeuArgTyrAsnArgGlnIleGlyGlnPhe 188  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 485 GAGGAAGCCAGATCAACAGCTCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTG 544  
OY 189 IleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGly 208  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 545 ATAGTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
OY 209 AlaValTyrLeuLeuAsnAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGln 228  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 605 GCTGTACCTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664  
OY 229 GluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSer 248  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 665 GAATTCACAGCCAGCTGCGGAGATTCCTGCGGCCAGCTGCGCTGCTGCTGCTGCT 724  
OY 249 GlyLeuLeuAlaLeuArgProGlySerSerLeuArgGlyLeuArgThrLeuProTyrAlaHis 268  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 725 GGGCTGTGGCCCTGCGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784  
OY 269 IleuValAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 785 CTCAGAGCTGCCCCCTTCTCCACCTTCTGAGCTTCTCCAGGTAC 832  
RESULT 3  
AAK56000  
ID AAK56000 standard; DNA; 1421 BP.  
XX  
AC AAK56000;  
XX  
DT 15-JUL-1999 (first entry)  
XX  
DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.  
XX  
KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 92..841  
FT /tag= a  
FT /product= "Apo-3 ligand"  
XX  
XX W09919490-A1.  
XX  
PD 22-APR-1999.  
XX  
PF 09-OCT-1998; 98WO-US21407.  
XX  
PR 17-DEC-1997; 97US-0069862.  
PR 10-OCT-1997; 97US-0062037.  
XX  
PA (GEMV ) GENENTECH INC.  
PI Ashkenazi AJ, Marsters SA, Pittl R;  
PI WPI: 1999-287982/24.  
DR P-PSDB: AAT09369.  
XX  
XX New human Apo-3 ligand (a tumour necrosis factor) homologue  
XX  
XX Claim 18; Fig 1; 74pp; English.  
XX  
XX The present sequence encodes a human tumour necrosis factor (TNF) and  
XX lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
XX cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
XX mammalian cancer cells, to induce NF-kappaB-dependent transcription and



Score: 1366.00 Matches: 268  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.60% Indels: 0  
DB: 21 Gaps: 0

US-09-245-198A-4 (1-284) x AAA49717 (1-1353)

```

OY 17 ArgSerLeuGlySerArgApelYglYalValArgInalagInProProAlaPmet 36
DB 1 CGATCCCTCGGGTCCCGGGATGGGGGGGTGAGGACAGCACGCCCCCGCCCATG 60
OY 37 AlaAlaArgSerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeu 56
DB 61 GCCGCCCTCGAGCCAGAGCGAGGGGGCCCGGGGGAGCCGGGACCCCGCTGTG 120
OY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76
DB 121 GTCCCGCTCGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 180
OY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96
DB 181 GTCACTTTGGGAGCGGGCATCGCTGCCCGCAGAGCCTGCCAGAGAGAGCTGGTG 240
OY 97 AlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro 116
DB 241 GCAGAGGAGAGACAGGACCCGTCGGAATCCCGACAGAGAGAGAGAGAGATCCT 300
OY 117 AlaProPheLeuAsnArgLeuValArgProArgSerAlaProGlySylArgLysThr 136
DB 301 GCGCTTCCTCAACCGACTAGTTCGGCTCCAGAGAGTGCACCTAAAGCGCGAGAAACA 360
OY 137 ArgAlaArgArgAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGly 156
DB 361 CGGGCTCGAAGAGCATGCGACCCCATTAATGATTCACAGACTGGAGCGGCA 420
OY 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrPglGluAlaArgIleAsnSerSer 176
DB 421 GCGGAGGAGAGTGTGGAGCGGAGCTGAGTGGTGGAGAGAGCCAGATCAACAGCTTC 480
OY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196
DB 481 AGCCCTCGCGCTACACCGCGAGATCGGGAGATTATAGTCACCGGCGCTGGCTCTAC 540
OY 197 TyrLeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeu 216
DB 541 TACCTGTACTGTGAGTGCATTTGATGAGGGAGCGTGTACTCTGAAGCTGGACTTG 600
OY 217 LeuValAlaAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleAsn 236
DB 601 CTGGTGATGGTGTCTGCTGGCCCTGGCGCTGCTGGAGAGAAATTCACGCCCTCGCGAGCT 660
OY 237 SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly 256
DB 661 TCCCTCGGAGCCCGACCTCGCTGCGCAGGTGTGGGCTGTGGCGCTGGCGGAGG 720
OY 257 SerSerLeuArgGlyLeuArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThr 276
DB 721 TCTCTCCGCGGATCCGACCTCCCTGGGCGCATTCAGGCTGGCCCTTCCTCAC 780
OY 277 TyrPheGlyLeuPheGlnValHis 284
DB 781 TACTTCGAGACTTTCAGAGTTCAC 804

```

RESULT 5  
ID ABR40255 standard; cDNA: 1353 BP.  
XX ABR40255;  
XX AC  
XX 15-JUN-2002 (first entry)  
XX DE cDNA encoding human PRO207 polypeptide.

XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;  
KW Leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;  
KW Inflammatory disorder; immune disorder; angiogenic disorder;  
KW gene therapy; cyostatic; neuroprotective; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200153486-A1.  
XX PD 26-JUL-2001.  
XX  
XX 11-FEB-2000; 2000WO-US03565.  
XX  
XX 08-MAR-1999; 99WO-US05028.  
XX PR 11-MAR-1999; 99US-123972P.  
XX PR 11-MAY-1999; 99US-133459P.  
XX PR 02-JUN-1999; 99WO-US12252.  
XX PR 22-JUN-1999; 99US-140650P.  
XX PR 22-JUN-1999; 99US-140653P.  
XX PR 20-JUL-1999; 99US-144758P.  
XX PR 26-JUL-1999; 99US-145698P.  
XX PR 28-JUL-1999; 99US-146222P.  
XX PR 17-AUG-1999; 99US-149395P.  
XX PR 31-AUG-1999; 99US-151689P.  
XX PR 01-SEP-1999; 99WO-US20111.  
XX PR 15-SEP-1999; 99WO-US21090.  
XX PR 30-NOV-1999; 99WO-US28313.  
XX PR 01-DEC-1999; 99WO-US28301.  
XX PR 01-DEC-1999; 99WO-US28334.  
XX PR 05-JAN-2000; 2000WO-US00219.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;  
XX Marsters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;  
XX Watanabe CK, Wood WI;  
XX  
XX WPI. 2002-205567/26.  
XX P-PSDB; AA086129.  
XX  
XX Thirty five nucleic acids encoding PRO polypeptides, useful for  
XX treating benign or malignant tumours, leukaemias and lymphoid  
XX malignancies, inflammatory, angiogenic and immunologic disorders -  
XX  
XX Claim 50; Fig 3; 302pp; English.  
XX  
XX The present invention relates to the isolation of novel human PRO  
XX polypeptides and the polynucleotide sequences encoding them. The  
XX PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
XX useful for treating benign or malignant tumours (e.g. renal, kidney,  
XX bladder, breast, etc), leukaemias and lymphoid malignancies, other  
XX disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
XX macrophagal, stromal and blastocoele disorders, inflammatory, immune  
XX and angiogenic disorders. The polynucleotide sequences are also  
XX useful in gene therapy. ABR40254-ABR40288 encode for the human PRO  
XX polypeptides of the invention.  
XX  
XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 3,966-90 Length: 1353  
XX Score: 1366.00 Matches: 268  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 94.60% Indels: 0  
XX DB: 24 Gaps: 0

US-09-245-198A-4 (1-284) x ABR40255 (1-1353)

```

OY 17 ArgSerLeuGlySerArgApelYglYalValArgInalagInProProAlaPmet 36
DB 1 CGATCCCTCGGGTCCCGGGATGGGGGGGTGAGGACAGCACGCCCCCGCCCATG 60

```

QY 37 AAlaAlaArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeu 56  
 Db 61 GCCGCCGCTGGAGCCAGAGAGGGGGCCGGGGGGAGCCGGGACCGCCCTGCTG 120  
 QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76  
 Db 121 GTCCCGCTGGCTGGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 180  
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
 Db 181 GTGAGTTGGGGAGCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 97 AlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspPro 116  
 Db 241 GCAGAGAGGAGCCAGAGACCCCTCGAAGCTGAAATCCAGACAGAAAGAAAGCAGAGATCCT 300  
 QY 117 AlaProLeuLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 136  
 Db 301 GCGCCCTTCTGAAACCGACTAGTTGCGGCTGCGAGAAAGTGCACCTAAAGCCGGAACA 360  
 QY 137 ArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly 156  
 Db 361 CCGGCTCGAAGAGGATCCGAGCCCATTTAATTCATCCACACCTGGACAGAGCGGA 420  
 QY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluAlaArgIleAsnSerSer 176  
 Db 421 GCGGAGCAGGTGAGAGCGGAGACAGTAGTGCTGGAGAGAACAGATCAACAGCTCC 480  
 QY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196  
 Db 481 AGCCCTTGCGCTACAAACCGCAGATCGGGAGTTTATGTCACCCGGGCTGCTTAC 540  
 QY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216  
 Db 541 TACTGTACGTGACGTGACATTTGATGAGGAGGAGGCTGCTCAACCTGAGCTGACTG 600  
 QY 217 LeuValAspGlyValIleAlaAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSer 236  
 Db 601 CTGGTGATGTGTGCTGGGCTGCTGCTGCTGAGAGAAATCTCAGCCACTGCGCGAGT 660  
 QY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 Db 661 TCCTCGGGGCGCCAGCTCCGCTGCTGCAAGTGTCTGGGCTGTGGCCCTGGCGCAGGG 720  
 QY 257 SerSerLeuArgIleArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThr 276  
 Db 721 TCCCTCGGAGATCCGACCCCTCCCTGGGCGCATCTCAAGGCTGGCCCTTCTCACC 780  
 QY 277 TyrPheGlyLeuPheGlnValHis 284  
 Db 781 TACTTCGAGCTCTTCCAGGTTCC 804  
 RESULT 6  
 AAV47613 standard; cDNA; 1236 BP.  
 ID AAV47613  
 AC AAV47613:  
 XX 27-OCT-1998 (first entry)  
 DE TNF related endothelium proliferative agent gene.  
 KW ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 1..750  
 FT /\*tag- a  
 FT /product- "TREPA"  
 XX

PN WO9835061-A2.  
 XX 13-AUG-1998.  
 PD 12-FEB-1998; 98NO-US02859.  
 PF 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 XX WALLEY SR.  
 XX  
 DR WPI: 1998-447255/38.  
 DR P-PSDB; AAW29745.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 PS Claim 11; Page 123-4; 142pp; English.  
 XX

CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 XX

SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Alignment Scores:  
 Pred. No.: 7,37e-83 Length: 1236  
 Score: 1265.00 Matches: 248  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.60% Mismatches: 0  
 Query Match: 87.60% Indels: 0  
 Db: 19 Gaps: 0

US-09-245-198A-4 (1-284) x AAV47613 (1-1236)

QY 36 MetaAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55  
 Db 1 ATGGCCGCCCGCTCGAGCCAGAAAGGAGGGGCGCGGGGGAGCCGGGACCGCCCTG 60  
 QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75  
 Db 61 CTGGTCCCGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 120  
 QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 95  
 Db 121 GTGGTCAAGTTGGGGAGCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAsp 115  
 Db 181 GTGGCAGAGAGGAGCCAGAGACCCCTCGAAGCTGAAATCCAGACAGAAAGAAAGCAGAGAT 240  
 QY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 135  
 Db 241 CCTGCGCTTCTGAAACCGACTAGTTGCGGCTGGAAGATGCACCTAAAGCCGGAAGA 300  
 QY 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155  
 Db 301 ACAGGGGCTCGAAGAGGAGATCGAGCCCATTTATGATTCACACAGACCTGGAGAGGAC 360



OY 156 GYAlAGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175  
 DB 361 GGAGCGCAGGAGGTGTGAGCGGACAGTGAAGTGGCTGGAGAGAACGAGATCAACAGC 420  
 OY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195  
 DB 421 TCCACCCCTCCGCTACACCGCCAGATCGGGAGATTATGATCACCAGGCGCTGGGCTC 480  
 OY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215  
 DB 481 TACTACCTGTAAGTGTGAGTGCATTTGATGAGGGGAGGAGGCTGTACCTGAAGCTGGAC 540  
 OY 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235  
 DB 541 TTGCGTGGAGATGAGTGTGCTGGCCCTGGCTGCGAGAAATTCACAGCCCTGGCGG 600  
 OY 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255  
 DB 601 AGTTCCTCGGGCCCGCCACTCCGCTCCAGGTGTGGGCTGTGGCCCTGGCGGCA 660  
 OY 256 GlySerSerLeuArgIleArgThrLeuProTrpAlaHisIleuLysAlaAlaProPheLeu 275  
 DB 661 GGGTCTCCCTGGGATCGGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTC 720  
 OY 276 ThrTyrPheGlyLeuPheGlnValHis 284  
 DB 721 ACCTACTTCGGACTCTCCAGGTTTCAC 747

## RESULT 7

AAD04350 ID AAD04350 standard; cDNA; 1236 BP.

AC AAD04350;

DT 04-JUL-2001 (first entry)

DE Human TREPA (TNF related endothelium proliferative agent) cDNA.

KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;

KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;

KW grafting; vulnery; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 1..750

FT /tag- a

FT /product- "Human TREPA (TNF related endothelium

FT proliferative agent)"

FT US6207642-B1.

PN 27-MAR-2001.

PD 26-JUN-1998; 98US-0105343.

PE 12-FEB-1997; 97US-0798692.

PR 10-FEB-1998; 98US-0021706.

XX (ABBO ) ABBOTT LAB.

PI Wiley SR;

XX WPI: 2001-280760/29.

DR P-PSDB; AAE00891.

XX Inducing angiogenesis in mammal at desired sites for promoting wound

XX healing, by administering soluble fragment of extracellular domain of

XX tumor necrosis factor related endothelium proliferative agent protein

XX Example 2; Column 73-74; 53pp; English.

CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present sequence is a cDNA clone ID #690050 encoding human TREPA.  
 XX

## SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:

## Alignment Scores:

Pred. No.: 7,37e-83 Length: 1236  
 Score: 1265.00 Matches: 248  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.60% Mismatches: 0  
 Query Match: 87.60% Indels: 0  
 DB: Gaps: 22

US-09-245-198a-4 (1-284) x AAD04350 (1-1236)

OY 36 MetAlaAlaArgArgSerGlnArgArgGlyArgGlyGluProGlyThrAlaLeu 55

DB 1 ATGGCCGCCCGTCGGAGCCAGAACCGGAGGGGCGCCGGGGGAGCCGGGCAACCGCCTG 60

OY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75

DB 61 CTGGTCCGCTGCGCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 120

OY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeu 95

DB 121 GTGGTACGTTGGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180

OY 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAsp 115

DB 181 GTGGCAGAGAGAGACAGACAGACCGCTCGGAGAACGATCCCGACAGAGAAAGCAGGAT 240

OY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135

DB 241 CCTGGCCCTTCTTCTTACCGAGTACTTGGCTCGCAGAGAGTGCACCTTAAGGCGGANA 300

OY 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155

DB 301 ACAGGGGCTCGAAGAGCGAGTGCAGCCCATTTATGAGTTCAATCCAGACAGCTGGACAG 360

OY 156 GYAlAGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175

DB 361 GGAGCGCAGGCGAGTGTGAGCGGACAGTGAAGTGGCTGGAGAGCCAGAAATCAACAGC 420

OY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195

DB 421 TCCACCCCTCCGCTACACCGCCAGATCGGGAGATTATGATCACCAGGCGCTGGCTC 480

OY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215

DB 481 TACTACCTGTAAGTGTGAGTGCATTTGATGAGGGGAGGCGCTGTACTGAGCTGGAC 540

OY 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235

DB 541 TTGCTGTGAGATGAGTGTGCTGGCCCTGGCTGCGCTGGAGAAATTCACAGCCACTCG 600

OY 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255

DB 601 AGTTCCTCGGGCCCGCCACTCCGCTCGCAGGTGTGGGCTGTGGCCCTGGCGGCA 660

OY 256 GlySerSerLeuArgIleArgThrLeuProTrpAlaHisIleuLysAlaAlaProPheLeu 275

DB 661 GGGTCTCCCTGGGATCGGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTC 720

OY 276 ThrTyrPheGlyLeuPheGlnValHis 284

DB 721 ACCTACTTCGGACTCTCCAGGTTTCAC 747

RESULT 8  
 AAS03964  
 ID AAS03964 standard; DNA; 898 BP.  
 XX  
 AC AAS03964;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.  
 XX  
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubecsis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 KW fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Location/Qualifiers  
 FT 52..873  
 FT CDS /tag- a  
 FT /product- "Fusion protein comprising a growth hormone  
 FT leader, a leucine zipper multimerisation  
 FT domain, and human TWEAK extracellular  
 FT domain"  
 XX  
 PN WO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PE 19-DEC-2000; 2000WO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMM) IMMUNEX CORP.  
 XX  
 PI Wiley SR.  
 DR WPI: 2001-417975/44.  
 DR P-PDB: AAO03459.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1; Page 39-40; 46pp; English.  
 XX  
 CC The sequence represents a DNA from the expression vector  
 CC pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rubecsis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and

CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 XX  
 SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1,326-68 Length: 898  
 Score: 1066.00 Matches: 207  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 73.82% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAS03964 (1-898)  
 QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAla 97  
 DB 250 AGTTTGGGAGCCGGGCGATCGCTGCCGCCAGAGCCCTGCCAGAGAGGCTGTGGCA 309  
 QY 98 GlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspProAla 117  
 DB 310 GAGGAGAGCCAGGACCCGTCGGAATGCCAGACAGAAAGAACCCAGGATCCTGCG 369  
 QY 118 ProPheLeuAsnArgLeuValArgProArgSerAlaProLysGlyArgLysThrArg 137  
 DB 370 CTTTCTGACCGACTAGTGTGGCCCTCCAGAGGACACCTAAAGCCGGAAACACGG 429  
 QY 138 AlaArgArgAlaIleAlaAlaHisTYrGluValHisProArgProGlyLysAspGlyAla 157  
 DB 430 GCTCGAAGAGCGATCGACGCCCATTTATGAATTCATCCACGACTGGACAGGACGAGCG 489  
 QY 158 GlnAlaGlyValAspGlyThrValSerGlyTTPGlnGluAlaArgIleAsnSerSerSer 177  
 DB 490 CAGGACAGGTGGACGAGGACAGTGAAGTGGCTGGAGAGAACCCAGATCAACAGCTCCACG 549  
 QY 178 ProLeuArgTYrAsnArgGlnIleGlyLysPheIleValThrArgAlaGlyLeuTYr 197  
 DB 550 CCTCGCGCTACACACGCCAGATCGGGAGTTATAGTCACCGCGGCTGGCTCTACTAC 609  
 QY 198 LeuTYrCysGlnValHisPheAspGluGlyLysAlaValTYrLeuLysLeuAspLeu 217  
 DB 610 CTGTACTGTACAGGTGACCTTTGATGAGGGAGAGCGCTGTACTGTAAGCTGACCTGTG 669  
 QY 218 ValAspGlyValLeuAlaLeuArgCysLeuGlnGluLysPheSerAlaThrAlaAspSer 237  
 DB 670 GTGATGCTGTGCTGGCCCTGGCTGCTCGAGAGATTTCTCAGCCACCTGGGCGAATTCC 729  
 QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySer 257  
 DB 730 CTCGGGCCCCAGCTCCGCTGCTCCAGAGTGTGGCTGTGGCCCTGGGCGCCAGCGTCC 789  
 QY 258 SerLeuArgIleArgThrLeuProTTPAlaHisLeuLysAlaAlaProPheLeuThrTYr 277  
 DB 790 TCCCTCGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCTACCTTAC 849  
 QY 278 PheGlyLeuPheGlnValHis 284  
 DB 850 TTGGACTCTTCCAGGTTCCAC 870  
 RESULT 9  
 AAX23424  
 ID AAX23424 standard; DNA; 1030 BP.  
 XX  
 AC AAX23424;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNRL3 DNA.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KW developmental abnormality; gestational abnormality; prostate cancer;

KW AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..627  
 FT CDS /tag= a  
 FT /product= "TNRL3"  
 XX  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNITV ) UNITV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR P-PSDB; AAW93590.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Example VII; Fig 13A; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: AP04, AP06, AP08 and AP09 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. AP04 is useful for diagnosing prostate cancer  
 CC by determining levels of AP04 in an individual. Prostate cancer can also  
 CC be treated using AP04 selective binding agents linked to a therapeutic  
 CC moiety. AP04 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. AP04 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in AP04  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using AP04 polypeptides/active  
 CC fragments and AP04 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of AP04 and detecting a change in level of AP04  
 CC activity. The method is performed in vivo or in vitro. AP0 polypeptides  
 CC are all useful as immunogens for preparing antibodies. AP04 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. AP08 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX  
 SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;  
 XX  
 XX  
 Alignment Scores:  
 Pred. No.: 3, 01e-68 Length: 1030  
 Score: 1062.00 Matches: 207  
 Percent Similarity: 99.52% Conservative: 0  
 Best Local Similarity: 99.52% Mismatches: 1  
 Query Match: 73.55% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAX23424 (1-1030)  
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
 DB 1 GTGAGTTTGGGAGCCGGGATGCTGTCGCCAGAGAGCTCCAGAGAGAGCTGCTG 60  
 QY 97 AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGlnSerGlnAspPro 116  
 DB 61 GCAG 120

QY 117 AlaProPheLeuAsnArgLeuValArgProArgSerAlaProLysGlyArgLysThr 136  
 DB 121 GCGCTTTCTTGAACCGACTAGTTCGGCTCCGAGAGAGTCACTTAAGGCCGGAACA 180  
 QY 137 ArgAlaArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 156  
 DB 181 CGGAGTCGAAG 240  
 QY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrPheGluGlnAlaArgLysSer 176  
 DB 241 GCGCAGCAGAGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 177 SerProLeuArgTyrAsnArgGlnIleGluPheIleValThrArgAlaGlyLeuTyr 196  
 DB 301 AGCCCTCTGCGCTACTACCGCCAGATCGGAGATTATATCTACCCGCGGCTGCTTAC 360  
 QY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 216  
 DB 361 TACCTGTACTGTACAGGTGACCTTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGlnPheSerAlaThrAlaIle 236  
 DB 421 CTGCTGATGTGTGTGCTGCGCTGCGCTGCGCTGAGAGATTCACAGCTGCGGCGAG 480  
 QY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 256  
 DB 481 TCCCTGGGGGCGAGCTCCGCTGTGCAAGTGTGTGGGTGTGGCTGCGGCGCAAGG 540  
 QY 257 SerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 276  
 DB 541 TCCCTCCCTGGAGATCCGCAACCTCCCTGGGCCCATCTCAAGCTGCCCCCTTCAC 600  
 QY 277 TyrPheGlyLeuPheGlnValHis 284  
 DB 601 TACTTGAGACTCTCCAGGTTTAC 624  
 RESULT 10  
 AA18599  
 ID AA18599 standard; cDNA; 1168 BP.  
 XX  
 AC AA18599;  
 XX  
 DT 21-JUL-1998 (first entry)  
 XX  
 XX Mus musculus tumour necrosis factor related ligand (TNRL) gene.  
 DE  
 XX  
 XX TNRL; tumour necrosis factor related ligand; tnfr; treatment;  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection; ds.  
 XX  
 XX Mus musculus.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT 2..679  
 FT CDS /tag= a  
 FT /note= "tumour necrosis factor related ligand"  
 XX  
 PN W09805783-A1.  
 XX  
 PD 12-FEB-1998.  
 XX  
 PF 07-AUG-1997; 97WO-US13945.  
 XX  
 PR 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 XX  
 PA (BIOL ) BIOGEN INC.  
 PA (UYGE-) UNITV GENEVA FACULTY MEDICINE.  
 PI Browning JL, Chicheportliche Y;  
 XX

DR WPI: 1998-145619/13.  
 DR P-PSDB; AAW47524.  
 XX Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 XX Claim 2; Pages 45-46; 69pp; English.

The sequence is that encoding mouse tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, auto-immune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon- gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. The DNA sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective amounts of a vector, e.g. a virus comprising a gene encoding TRELL. It may also be of use in the preparation of prepare probes for screening natural/synthetic DNAs for TRELL-encoding sequences and for antisense therapy.

SO Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;

Alignment Scores:  
 Pred. No.: 3,8e-65 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 DB: Gaps: 0

US-09-245-198A-4 (1-284) x AAV18599 (1-1168)

QY 61 LeuGLyLeuGLyLeuAlaLeuAlaAcylLeuGLyLeuLeuAlaValValSerLeuGLy 80  
 DB 5 CTGACCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 64  
 QY 81 SerArgAlaSerLeuSerAlaGlnLupProAlaGlnLupLysLeuValAlaLupLup 100  
 DB 65 AGCTGGGACAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124  
 QY 101 GlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspProAlaProPheLeu 120  
 DB 125 CGGAGCGCCCTGAACCTGAATCCCAAGACAGAAACCCAGATGTGCTCTTCTTG 184  
 QY 121 AsnArgLeuValArgProArgSerAlaProLysGlyArgLysThrArgAlaArgArg 140  
 DB 185 GAAACACTAGTCGGCCCTCGAAGAGTCTCTTAAGCCGGAAGCGCGGCTGCGCGA 244  
 QY 141 AlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
 DB 245 GCTATTGACGCCATTAGAGGTTTCATCTCGCCAGACAGATGAGACACAGCAGGT 304  
 QY 161 ValAspGlyThrValSerGlyTyrGlnGlnLupLupLysLeuAlaValValSerLeu 180  
 DB 305 GTGGATGGACAGTACAGTGGCTGGAGACAGCAAAATCAACAGCTCCAGCCCTGCGC 364  
 QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
 DB 365 TACACACCGCCAGATTGGGGAATTACAGTATCATGGGTGGCTGCTACTACTACT 424  
 QY 201 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
 DB 425 CAGGTGACCTTGTATGAGGGAAGCTGTCTCACTGAAGCTGAGCTGTGTGACAGGT 484  
 QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerLeuGlyPro 240

DB 485 GTGCTGGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544  
 QY 241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProLysSerLeuArg 260  
 DB 545 CACTCCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
 QY 261 IleArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280  
 DB 605 ATCCGACACCTCCCTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664  
 QY 281 PheGlnValHis 284  
 DB 665 TTTCAAGTTCAC 676  
 RESULT 11  
 AAX23425  
 ID AAX23425 standard; DNA; 701 BP.  
 XX AAX23425;  
 AC AAX23425;  
 DE 18-JUN-1999 (first entry)  
 DE Mouse TNRL3 DNA.  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 XX  
 XX Mus sp.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 1..636  
 FT /tag- a  
 FT /product- "TNRL3"  
 FT  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PE 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIM ) UNIV WASHINGTON.  
 PL  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR P-PSDB; AAW93591.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PS  
 PS Example VII; Fig 13B; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) family receptor polypeptides: APO4, APO6, APO8 and APO9 or their active fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or their active fragments. APO4 is useful for diagnosing prostate cancer by determining levels of APO4 in an individual. Prostate cancer can also be treated using APO4 selective binding agents linked to a therapeutic moiety. APO4 polypeptides are also useful for identifying selective binding agents, useful in diagnosis/treatment of disease by binding of agents to the polypeptide/active fragment which is extracellular, or expressed on the cell surface. The binding is preferably performed in vivo. APO4 polypeptides/active fragments are also useful for screening for agonists and antagonists by binding and observing the change in APO4 activity. Effective pharmacological agents useful in diagnosis or treatment of disease are also identified using APO4 polypeptides/active

CC fragments and ApO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4  
 CC activity. The method is performed in vivo or in vitro. Apo polypeptides  
 CC are all useful as immunogens for preparing antibodies. ApO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.

SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,25e-61	Length:	701
Score:	968.00	Matches:	188
Percent Similarity:	93.36%	Conservative:	9
Best Local Similarity:	89.10%	Mismatches:	14
Query Match:	67.04%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-4 (1-284) x AAV24140 (1-701)

```

OY 74 LeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu
DB 1 CTGGTCGTGGTACGCTGGGAGCTGGCAACGCTGTCTGCCAGAGCCTTCTCAGAG
OY 94 GluLeuValAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu
DB 61 GAGCTGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
OY 114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGly
DB 121 CAGGATGTGGTACCTTTCTTGGACAACTAGTCGCGCTCGAAGAAAGTCTCTAAAGC
OY 134 ArgLysThrArgAlaArgArgAlaLeuAlaAlaHisTyrGlnValHisProArgProGly
DB 181 CGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
OY 154 GlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrGlnGluGluAlaArg
DB 241 CAGGATGGAGCACAAGCAGCTGTGATGGACAGTGCAGTGCAGTGCAGTGCAGTGCAG
OY 174 AsnSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAla
DB 301 AACACCTCCAGCCCTCTGCGTACGACAGCCGCGAGATGGGGAATTACAGTCATCAGG
OY 194 GlysLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeu
DB 361 GGGCTCTACTACTCTGTACTGTAGGTGGCACTTGTGATGAGGAAAGGCTGTCTAAG
OY 214 LeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThr
DB 421 CTGGACTGTGCTGTGCAACAGGTGTCTGCGCCCTGCGCTGCGAAGATTTCTCAGCC
OY 234 AlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeu
DB 481 GCAGCAAGCTCTCTGCGGCGCCAGCTCCGTTGTGTCAGAGTGTCTGAGGCGTGG
OY 254 ArgProGlySerSerLeuArgGlyLeuArgThrLeuProTyrAlaHisLeuLysAlaAlaPro
DB 541 CGGCGAGGCTCTCCCTTGGATCCGACACCTCCCTGGGCTCATTAAGGCTGCCCC
OY 274 PheLeuThrTyrPheGlyLeuPheGlnValHis 284
DB 601 TTCCCTAACCTACTTTGAGCTCTTTCAAGTTTCA 633

```

#### RESULT 12

AAV24140 ID AAV24140 standard; DNA; 2856 BP.

AAV24140;

AC AAV24140;  
 XX 28-SEP-1998 (first entry)  
 XX Homo sapiens BARD1 exon 1.

XX BARD1; ring protein; BRCA1; breast cancer; risk; diagnosis; exon; ss.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 XX Homo sapiens.  
 XX Homo sapiens.  
 FT Key location/Qualifiers  
 FT exon 2031..2188  
 FT /tag= a  
 FT /number= 1

W09812327-A2.

26-MAR-1998.

19-SEP-1997; 97WO-US16842.

04-APR-1997; 97US-0042985.

20-SEP-1996; 96US-0025296.

03-APR-1997; 97US-0042611.

(TEXA ) UNIV TEXAS SYSTEM.

Baer R, Bowcock AM;

WPI: 1998-230317/20.

DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which  
 as breast cancer antigen, BRCA1, binding proteins are useful to  
 identify patient having or at risk of developing cancer

Claim 25; Page 307-308; 348pp; English.

The sequence is that of containing a BARD1 exon which can be used in the  
 preparation of the recombinant breast cancer antigen, BRCA1, binding  
 proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the  
 detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid  
 sequence, specifically a wild type BARD1 composition for the detection  
 or purification of BRCA1, useful to identify a patient having, or at  
 risk of developing cancer. BARD1 can be used in the preparation of an  
 anti-BARD1 antibody, and in the detection and purification of a BRCA1  
 protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the  
 identification of a binding protein agonist or antagonist that alters  
 the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or  
 biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445  
 complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,  
 BE31 or BE445, a specific anti-BARD1 antibody can be used to identify  
 a patient having or at risk of developing cancer.

Sequence 2856 BP; 616 A; 778 C; 786 G; 674 T; 2 other;

#### Alignment Scores:

Pred. No.:	4.74	Length:	2856
Score:	127.00	Matches:	84
Percent Similarity:	36.00%	Conservative:	33
Best Local Similarity:	25.85%	Mismatches:	122
Query Match:	8.80%	Indels:	86
DB:	19	Gaps:	13

US-09-245-198a-4 (1-284) x AAV24140 (1-2856)

```

OY 14 ProLeuProArgSerLeuGlySerArgAspGlyAlaVal-----
DB 1788 CCCACTCGGCGCAGCGGCTGCTGTGCGGCGCCGCCCTGTGCGAGTCCCTATT 1847
OY 28 -----ArgGlnAlaGlnPro-----ProAlaProMetAlaAla 38
DB 1848 TGGAGCATTTGGCGGCGCGTCCGCCGCCGCCGCCGCCGCCGCCCTGTGCGGCGC 1907
OY 39 ArgArgSerGlnArgArgArgGlyArgArgGlyGluProGlyTyrAlaLeuValPro 58
DB 1908 CGCCGCTCCACGACGCGGAGCTTGGCGCGTGTGAGTCCGAGCTGCGAGCTTCCCT 1967
OY 59 LeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLacysLeuGlyLeuLeuLeuAlaValSer 78

```

	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

OY 131 ProlysglyAlaThrArgAlaArgAlaIleAlaHisIleValHisPro 150  
 DB 742 -----GGCGTGGCGGCGGAGCAGCATCGAGC-----CCT 771  
 OY 151 ArgProGlyGln-----AspGlyAlaGln-----158  
 DB 772 CGTCTGTCGTGGCGGCGGAGTCCCGGGGTTCTCGAGCGCGTGGCAACTTCGCGTC 831  
 OY 159 -----AlaGlyValAspGlyThrValSer 166  
 DB 832 CGGAGCAGCATCTCTACGTGCCAGCAGCAGCGCCCTGCGCCAGCGAGCATGTGAG 891  
 OY 167 -GlyTPGluGluAlaArgIleAsnSerSerProLeuArgThrValSerGlnIleG 186  
 DB 892 GAGCTTGAGAGCGCCATCTCGAAGTCCCGTCCCGCGCGTTCGGCGCCGTCAGAGC 951  
 OY 186 ygluPheIleValThrArgAlaGlyLeuTyrrTyrrTyrrGlyValHisPheAspG 206  
 DB 952 GCCGCTGCTGTCACGAGCAGCATGAGCGCTTCCTCGCGCGCTCGCGTCCGCGAGC 982  
 OY 206 uGlyLysAlaValTyrrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArg 226  
 DB 983 -----TCCGCTGCGCGAGCAGCATGAGCGCTTCCTCGCGCGCTCGCGTCCGCGAGC 1038  
 OY 226 slauGluGluPheSerAlaThrAlaAlaSerSerLeuGly-----ProGlnLe 242  
 DB 1039 CGCGGAGCCCTTCAGAGAGCGGCGCGTCCCAAGATGGCGAGTTCGCGCGCGAGTGT 1098  
 OY 242 uArgLeuGlyGlnValSerGlyLeuAlaLeuAlaArgProGlySerSerLeuArgIleAr 262  
 DB 1099 GCCCATGACGATCATCAGCGGGGTC-----GGCAACAGAGCGCGCGTCCG 1143  
 OY 262 gThLeuProThrAlaHis---LeuLysAlaAlaProPheLeuThrTyrrPheGly 279  
 DB 1144 GCTGTGTCTGTGGGCGAAGACTTCGAGCGCGCGCGAGGTGGCGCGCGG 1198

RESULT 14  
 AAV69289  
 ID AAV69289 standard; cDNA; 2832 BP.  
 AC AAV69289;  
 DT 01-MAR-1999 (first entry)  
 XX  
 DE Murine JNK-interacting protein 1 (JIP-1) cDNA.  
 KW JIP-1; JNK-interacting protein; c-Jun NH2-terminal kinase;  
 KW signal transduction; inhibitor; mouse; neurodegenerative disease;  
 KW Parkinson's disease; Alzheimer's disease; blood clot; stroke;  
 KW malignancy; cancer; leukemia; autoimmune disease; inflammation;  
 KW apoptosis; therapy; diagnosis; ds.  
 XX  
 OS Mus sp.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 180..2162  
 FT /\*\*tag= a  
 XX  
 XX WO9849188-A1.  
 XX  
 XX 05-NOV-1998.  
 XX  
 XX 28-APR-1998; 98WO-US08513.  
 XX  
 XX 28-APR-1997; 97US-0819177.  
 XX  
 XX (UYMA-) UNIV MASSACHUSETTS.  
 XX  
 XX Davis RJ, Dickens M;  
 XX  
 XX WPI; 1999-024042/02.  
 XX  
 XX P-PSDB; AAM81525.

XX  
 PT c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to  
 PT treat neurodegenerative disease, blood clot, leukemia, autoimmune  
 PT disease, and inflammation  
 XX  
 PS Claim 11; Page 66-70; 95pp; English.  
 XX  
 CC This cDNA sequence codes for murine JNK-interacting protein 1  
 CC (JIP-1, see AAM81625), a novel cytoplasmic anchor protein that  
 CC specifically binds to and inhibits the biological effects of JNK  
 CC (c-Jun NH2-terminal kinase), including the initiation of apoptosis  
 CC and oncogenic transformation. To isolate the clone, a yeast  
 CC two-hybrid method was used to screen a mouse embryo cDNA library to  
 CC identify proteins that interact with JNK. 7 overlapping partial  
 CC clones were isolated, and the full-length cDNA was subsequently  
 CC obtained by screening a mouse brain lambda ZAPII cDNA library. The  
 CC invention provides JIP-1 nucleic acids and polypeptides; expression  
 CC vectors and host cells. The JIP-1 polypeptide and nucleic acids  
 CC (including antisense and ribozyme molecules) can be used in the  
 CC manufacture of a medicament for treating a pathological condition  
 CC associated with abnormal expression or activity of JNK, such as a  
 CC neurodegenerative disease (selected from Parkinson's disease and  
 CC Alzheimer's disease), a blood clot, stroke, malignancy, leukemia,  
 CC an autoimmune disease or inflammation (all claimed).  
 XX  
 SQ Sequence 2832 BP; 569 A; 888 C; 840 G; 535 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.12 Length: 2832  
 Score: 124.50 Matches: 72  
 Percent Similarity: 36.06% Conservative: 25  
 Best Local Similarity: 26.77% Mismatches: 73  
 Query Match: 8.62% Indels: 100  
 DB: 20 Gaps: 12  
 US-09-245-198a-4 (1-284) x AAV69289 (1-2832)  
 OY 1 MetSerLeuLeuAspPheGluIleSerAlaArgLeuProLeuProArgSerLeuGly 20  
 DB 20 ATAGCTTGATATGCTCTCCGAGCCGCG-----CCGGCCGACGTCAGTCCGAA 70  
 OY 21 SerArgAspGlyGlyAlaValArgGlnAla-----30  
 DB 71 CCCCGCGCGCGCGCGCGCTCCACCGCTCCGCTGTCGCGCGCGCGCGCGCGCG 130  
 OY 31 GlnProProlaPromeAlaAlaArgArgSerGln-----ArgArgArgGlyArg 47  
 DB 131 CCGCTCCCGCGCGCGCTCTCCGCGGATGCGCAGGGCTGCACCCCGGAATGGCGAGCG 190  
 OY 48 ArgGlyGluProGlyThrAlaLeuLeuValProLeuAla-----LeuGlyLe 63  
 DB 191 AGAGAGCGGCTGTGGCGGGGCGCGCGCTCCACCGGCGCGCTCCCATTCCTGGACT 250  
 OY 63 u-----63  
 DB 251 GCACATCGCTGCGCTCCCAATTCAGGCTCACCCATGACATCAGCTGAGAGATTGA 310  
 OY 64 -----GlyLeuAlaLeuAlaCys-----69  
 DB 311 GGATGAAGACCTTTCGAGATCACTGACGAGTGTGATCAGCTGACGCTCAAGACAC 370  
 OY 70 -LeuGlyLeu-----LeuLeuAlaValValSerLeuGlySerArgAl 83  
 DB 371 CTTGTCTCTCGGCGCGCGCGCGCGCGCTGTCTCTCGGGTACGACGCGCGCGG 430  
 OY 83 aSerLeuSerAlaGlnGluProAlaGlnGlnLeuVal-----96  
 DB 431 GAGCGCGCTGACGCGGAGATGCTCAGATGAGACTGTCGACGCGGAGTGCACATCC 490  
 OY 97 -----AlaGluGluAspGlnAsp-----ProSerG 105  
 DB 491 GGGCGCGAGGACGACGAGGAGGAGAGAGAGAGAGAGTCTGCTGCCAAGCAGGAGT 550







Fri May 9 16:16:17 2003

us-09-245-198a-4.rng

Page 15

Oy	224	uagggysleuengunglupheserAlaThrAlaAserSerleuGlYProGlnleuArgle	244
Db	31408	ccgg-----TTCTGCGGGGGGGGGTTCG-----GGACCC-----	31378
Oy	244	ucysGlnValSerGlyleuLeuAlaLeuArgProGlySerSerleuArgIleArgThrle	264
Db	31377	-----GGCCTTGTGCGCTGACGGGTGCTTTGTTCACCGCGCTCACCCACCC	31331
Oy	264	uProTPrAlaHisleuLysAlaAlaPro	273
Db	31330	CCCCCACTCATACACCGCGCTGGGCC	31303

Search completed: May 8, 2003, 03:43:17  
Job time : 248.35 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:12 ; Search time 2227.92 Seconds

(without alignments)  
3709.824 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRPLPRSLG.....PMAHKAAPPLYTFGLFQVH 284

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp

-Q/cg2\_1/USPTO.spool/US09245198/runat\_06052003\_170239\_9615/app\_query\_fasta\_1.846

-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45

-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTPRM=psio -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USFR=US09245198 -ECGN\_1\_1\_3073 -runat\_06052003\_170239\_9615 -NCPV=6 -ICPU=3

-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_dat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_ov:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_in:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rdd:\*  
36: em\_htg\_rdd:\*  
37: em\_htg\_vtc:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1366	94.6	1353	6 AX201324	AX201324 Sequence
2	1366	94.6	1368	9 AF055872	AF055872 Homo sapi
3	1331.5	92.2	1651	9 BC019047	BC019047 Homo sapi
4	1298	89.9	1306	9 AF030099	AF030099 Homo sapi
5	1265	87.6	1236	6 AR140407	AR140407 Sequence
6	1066	73.8	898	6 AX180714	AX180714 Sequence
7	1020	70.6	1168	10 AF030100	AF030100 Mus muscu
8	589.5	40.8	215795	2 AC126876	AC126876 Homo sapi
9	589.5	38.6	177555	2 AC127470	AC127470 Pan trogl
10	557	30.8	177555	2 AC130192	AC130192 Sus scrof
11	555.5	38.5	161428	2 AC126925	AC126925 Canis fam
12	555.5	38.5	203083	2 AC069459	AC069459 Mus muscu
13	555.5	38.5	234182	10 AL603707	AL603707 Mouse DNA
14	495	34.3	153553	2 AC126921	AC126921 Bos tauru
15	489.5	33.9	138792	2 AC119115	AC119115 Rattus no
16	385.5	26.7	184026	2 AC098923	AC098923 Rattus no
17	214	14.8	203281	2 AC126237	AC126237 Canis fam
18	145.5	10.1	187835	8 AP002482	AP002482 Oryza sat
19	140	9.7	42210	1 SC1C2	AL031124 Streptomy
20	138.5	9.6	190050	1 AL646059	AL646059 Rattus no
21	136.5	9.5	242930	1 AC105470	AC105470 Rattus no
22	135	9.3	6397	1 AF002222	AF002222 Pseudomon
23	135	9.3	10728	1 AE004651	AE004651 Pseudomon
24	134	9.3	10965	1 AE004740	AE004740 Pseudomon
25	130	8.9	40352	1 SCC61A	AL356595 Streptomy
26	129	8.9	5751	1 SGGNASTRS	X89010 S. glaucosce
27	129	8.9	16185	1 SCC46	AL139164 Streptomy
28	129	8.9	25459	1 SGAJ5985	AT006985 Streptomy
29	129	8.9	32941	1 SCMI0	AL133469 Streptomy
30	129	8.9	43961	1 SCMI1	AL133422 Streptomy
31	128	8.9	137833	9 CNS01DV6	AL133455 Human chr
32	127	8.8	194304	9 AC016708	AC016708 Homo sapi
33	126.5	8.8	207050	1 AL646063	AL646063 Rattus no
34	125.5	8.7	999	9 HSBARD11	AF038034 Homo sapi
35	125.5	8.7	28795	1 AVINIEC	M20568 A. vinelandi
36	125.5	8.7	87144	8 AP003822	AP003822 Oryza sat
37	125.5	8.7	153595	2 AP004272	AP004272 Oryza sat
38	125	8.7	30590	1 SC4G2	AL031371 Streptomy
39	125	8.7	191221	2 AP005292	AP005292 Oryza sat
40	124.5	8.6	36734	1 SC6610	AL049497 Streptomy
41	124.5	8.6	98504	2 AP003876	AP003876 Oryza sat
42	124.5	8.6	107132	2 AP003911	AP003911 Oryza sat
43	124.5	8.6	154746	14 HSY2HG52	28609 Herpes simp
44	124.5	8.6	176580	2 AP003634	AP003634 Oryza sat
45	124	8.6	3957	6 A45258	A45258 Sequence 2

RESULT 1

# ALIGNMENTS

AX201324  
 LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001  
 DEFINITION Sequence 3 from Patent WO0153486.  
 ACCESSION AX201324  
 VERSION AX201324.1 GI:15391154  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1353)  
 AUTHORS Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L.,  
 Hillan, K.J., Masters, S.A., Pan, J., Pitt, R.M., Roy, M.A., Smith, V.,  
 Stone, D.M., Watanabe, C.K. and Wood, W.I.  
 TITL Positions and methods for the treatment of tumour  
 JOURNAL Patent: WO 0153486-A 3 26-JUL-2001.  
 Genentech, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..1353  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 257 a 443 c 389 g 264 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,66e-63 Length: 1353  
 Score: 1366.00 Matches: 268  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.60% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-09-245-198a-4 (1-284) x AX201324 (1-1353)  
 Oy 17 ArgSerLeuGlySerArgAspGlyGlyAlaValArgGlnAlaGlnProProAlaPromet 36  
 |||||||  
 Db 1 CGATCCCTCGGTCCTCCGGGATGGGGGGCGGTAGGCGACAGACCCCGCCCATG 60  
 Oy 37 AlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGlnProGlyThrAlaLeuLeu 56  
 |||||||  
 Db 61 GCCGCCCGTCGAGCAGAGCGAGGGGGCGCCGGGGGAGCGCGGACCGCCCTGCTG 120  
 Oy 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLeuGlyLeuLeuLeuAlaVal 76  
 |||||||  
 Db 121 GTCCCGCTCGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 180  
 Oy 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGlnGlnVal 96  
 |||||||  
 Db 181 GTACGTTTGGGGAGCCGGGCGATCTGCTCCGCCAGAGAGCTCTCCAGAGAGCTG 240  
 Oy 97 AlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspPro 116  
 |||||||  
 Db 241 GCAGAGAGAGACAGAGACCGCTCGGAATGATCCCGAGAGAGAGAGAGAGAGATCT 300  
 Oy 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyGlyArgGlyThr 136  
 |||||||  
 Db 301 GCCCTCTTCTGTAACGAGTCTGCGCTCGCAGAGAGTCCACCTAAAGGCCGGAACA 360  
 Oy 137 ArgAlaArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 156  
 |||||||  
 Db 361 CGGGCTCGAGAGACGATCCGACCTATGAAAGTTTCAATCCAGACCTGGACAGAGAGGA 420  
 Oy 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrGlnGlnGlnAlaArgIleAsnSer 176  
 |||||||  
 Db 421 GCGCAGAGAGGTGTGACGCGGACAGTGAAGTGTGGAGAGAGAGAGAGAGAGAGAGAG 480  
 Oy 177 SerProLeuArgTyrAsnArgGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyr 196  
 |||||||  
 Db 481 AGCCCTCTGCGTACAAACCGCATCGGAGAGTATATGATCACCGGCGTGGCTCTAC 540  
 Oy 197 TyrLeuTyrCysGlnValHisPheAspGlnGlyValAlaValTyrLeuLeuAspLeu 216  
 |||||||  
 Db 541 TACCTTACTGTCAGGTGACCTTGATGAGGGGAGAGGCTGTCTACCTGAAGCTGACTTG 600

Oy 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGlnPheSerAlaThrAlaAlaSer 236  
 |||||||  
 Db 601 CTGTGTGATGTCTGTCTGGCCCTGCGCTGGAGAGATCTCAGACCATGGGGAGAG 660  
 Oy 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 |||||||  
 Db 661 TCCCTGGGGCCAGCTCCCGCTCTGACAGGTCTGGCGCTGTGGCCCTGGCGCCAGG 720  
 Oy 257 SerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 276  
 |||||||  
 Db 721 TCTCTCCGCGGATCCGACCTCTCCCTGGGCCCATCTCAAGGCTGCCCTCTCTCAC 780  
 Oy 277 TyrPheGlyLeuPheGlnValHis 284  
 |||||||  
 Db 781 TACTTGGAATCTTCCAGGTTCAC 804  
 RESULT 2  
 AF055872 1368 bp mRNA linear PRI 04-MAY-1998  
 LOCUS AF055872  
 DEFINITION Homo sapiens Ap03/DR3 ligand (APO3L) mRNA, complete cds.  
 ACCESSION AF055872  
 VERSION AF055872.1 GI:3108230  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1368)  
 AUTHORS Masters, S.A., Sheridan, J.P., Pitt, R.M., Brush, J., Goddard, A. and  
 Ashkenazi, A.  
 TITL Identification of a ligand for the death-domain-containing receptor  
 JOURNAL Curr. Biol. 8 (9), 525-528 (1998)  
 MEDLINE 98228355  
 PUBMED 9560343  
 REFERENCE  
 2 (bases 1 to 1368)  
 AUTHORS Masters, S.A., Sheridan, J.P., Pitt, R.M., Brush, J., Goddard, A. and  
 Ashkenazi, A.  
 TITL Direct Submission  
 JOURNAL Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,  
 South San Francisco, CA 94080, USA  
 FEATURES  
 source location/Qualifiers  
 1..1368  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17p13"  
 1..1368  
 /gene="APO3L"  
 58..807  
 /function="binds to the death-domain-containing receptor  
 Ap03/DR3"  
 /note="TNF homolog"  
 /codon\_start=1  
 /product="Ap03/DR3 ligand"  
 /protein\_id="AAC39724.1"  
 /db\_xref="GI:3108231"  
 /translation="MAARSSORRRRRRGGPCALLVPLALGIGLALACGLLAVSL  
 GSRASLSQEPQAEELVAEEDDPSELNQTESODPAFLNRLVPRRSAPKGRTR  
 ARRAIAAHVEVHPRGODGAQGVGTSGWEARINSSPLRNROIQEFTVTRAGL  
 YLYCOVHDEKAVYLRKLDLVGVLRRCLEERSAVTAAASLGPOLRLGVSGLLAT  
 RGGSLRTITLPMALHKAPFLTYGLOVH"  
 BASE COUNT 272 a 443 c 389 g 264 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,71e-63 Length: 1368  
 Score: 1366.00 Matches: 268  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.60% Indels: 0

DB: 9 Gaps: 0

US-09-245-198a-4 (1-284) x AF055872 (1-1368)

QY 17 ArgSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProProAlaProMet 36  
 1 CGATCCCTCGGGTCCCGGATGGGGGGGGGAGGAGCAGCAGCCCGCCCGCCAG 60

QY 37 AlaAlaArgArgSerGlnArgArgArgGlyArgGlyGlnProGlyThrAlaLeu 56  
 61 GCGCCCGGTCGAGCAGAGGGGAGGGGCGGGGGGAGCCCGGCGACCGCCGCTG 120

QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76  
 121 GTCGCCCTCGCCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 180

QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGlnLeuVal 96  
 181 GTCAGTTTGGGAGCGGGGAGTCGCTGTCGCCAGAGCCTGCGCAGAGAGCTG 240

QY 97 AlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspPro 116  
 241 GAGAGAGAGACAGACCGCGTCGGAACGAAATCCCGACAGAGAGAGAGAGAG 300

QY 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThr 136  
 301 GCGCCCTTCTGACCGAGTACTGCGGCTCCGAGAGAGTCCACTAAAGCGCGA 360

QY 137 ArgAlaArgArgAlaLeuAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGly 156  
 361 CCGGCGCAGAGAGCGATCCAGCCCATATGAAATTCATCCAGCAGCTGAGCAG 420

QY 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrGlnGlnAlaArgLysSerSer 176  
 421 GCGAGGCGAGGTGGAGCGAGAGTGAAGTGGTGGAGAGAGCCAGAAATCAAC 480

QY 177 SerProLeuArgTyrAsnArgGlnLeuGlyLysPheLeuValThrArgAlaGlyLeuTyr 196  
 481 AGCCCTCTCGGTACACAGCGGAGATGGGAGTTATATAGTACCGGCGCTGG 540

QY 197 TyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeu 216  
 541 TACCTGTACTGTACAGTGCATTTGATGAGGGAGAGCTGTCTACCTGAAG 600

QY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGlnPheSerAlaThrAlaAlaSer 236  
 601 CTGGTGATGTGTGCTGCGCTGCGCTGCTGGAGAAATCTCCAGCCACTGG 660

QY 237 SerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 661 TCCCTCGGCGCCAGCTCGGCTCTGCGAGTGTCTGGGCTGTGGCCCTGCG 720

QY 257 SerSerLeuArgTyrLeuArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThr 276  
 721 TCCCTCCCTCGGATCCGACACCTCCCTGCGGCAATCAAGGCTGCCCTCT 780

QY 277 TyrPheGlyLeuPheGlnValHis 284  
 781 TACTTGGACTCTTCAGGTTTAC 804

RESULT 3  
 BC019047 1651 bp mRNA linear PRI 11-DEC-2001  
 LOCUS Homo sapiens, Similar to tumor necrosis factor (ligand)  
 DEFINITION superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,  
 complete cds.  
 ACCESSION BC019047  
 VERSION BC019047.1 GI:17512138  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1651)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Louis Staudt  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, D.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

FEATURES  
 source 1. 1651  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20669 IMAGE:4766071"  
 /tissue\_type="Primary B-cells from Tonsils"  
 /clone\_id="NH.MGC-48"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 106..510  
 /codon\_start=1  
 /product="Similar to tumor necrosis factor (ligand)  
 superfamily, member 12"  
 /protein\_id="AAH19047.1"  
 /db\_xref="gi:17512138"  
 /translation="MAARRSQRRRRRREPEPTALLVPLALGIALACGLILAVSL  
 GSRASLSQEPAEELVAEDDPSLELNPQTEESODPAPFLNRLVRPRRSPKGRKTR  
 ARRAIAHYEVRPGODGAGADGGVTCILRP"

CDS  
 106..510  
 /product="Similar to tumor necrosis factor (ligand)  
 superfamily, member 12"  
 /protein\_id="AAH19047.1"  
 /db\_xref="gi:17512138"

BASE COUNT 344 a 517 c 481 g 309 t

ORIGIN

Alignment Scores:  
 Pred. No.: 3,53e-61 Length: 1651  
 Score: 1331.50 Matches: 275  
 Percent Similarity: 77.25% Conservative: 0  
 Best Local Similarity: 77.25% Mismatches: 1  
 Query Match: 92.21% Indels: 81  
 DB: 9 Gaps: 1

US-09-245-198a-4 (1-284) x BC019047 (1-1651)

QY 9 SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg 28  
 DB 25 TCCGCGCGCGGCTCCCTCCCTCCCGATCCCTCGGCTCCCGGATGGGGGGGAG 84

QY 29 GlnAlaGlnProProAlaProMetAlaAlaArgArgSerGlnArgArgArgArgArg 48  
 DB 85 CAGGACAGACCCCGCCCGCCCGATGCGCGCGCGCGAGCAGAGGCGGAGGCGCC 144

QY 49 GlyLysProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaVal 68  
 DB 145 GGGAGACCGGGGACCGCCCTGCTGTCGCGCTGGGCTGGGCTGGGCTGGGCTGG 204



```

|||||
Db 423 GCCAGATACACGCTCCAGCCCTCGCCTACACCCGACATCGGGAGTTTATAGTC 482
Oy 191 ThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluIlylsAlaVal 210
Db 483 ACCCGGGCTGGCTACTACTGTACTGTACAGTGTACTTGTATGATGAGGGAAGGCTGTC 542
Oy 211 TyrIleuIysLeuAspIleuValAspGlyValIleuAlaLeuArgCysLeuGluGluPhe 230
Db 543 TACCTGAAGCTGACTGCTGTGGAGTGGTGGCTGGCCCTCGCTGCGTGGAGGAATTC 602
Oy 231 SerAlaThrAlaIleSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeu 250
Db 603 TCAGGCACATGGCGGACGTTCCCTCGGGGCCACAGCTCCGCTCTGCCAGGTCTGTGGCTG 662
Oy 251 LeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLys 270
Db 663 TTGGCCCTGCGGCGCAGGGTCTCCCTGCGGATCCGACCCCTCCGCGCTCAAG 722
Oy 271 AlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
Db 723 GCTGCCCTTCCTCACTTCTGAGATCTTCACAGTTTCAC 764

RESULT 5
AR140407 1236 bp DNA linear PAT 16-JUN-2001
LOCUS AR140407 Sequence 1 from patent US 6207642.
ACCESSION AR140407
VERSION AR140407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Wiley,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
disease
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
source 1.1236
location/Qualifiers
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 7.51e-58 Length: 1236
Score: 1265.00 Matches: 248
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 87.60% Indels: 0
Gaps: 0
DB: 6

US-09-245-198a-4 (1-284) x AR140407 (1-1236)
Oy 36 MetAlaAlaArgSerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeu 55
Db 1 ATGGCGCGCCGTCGAGGACGAGAGCGGAGGCGCGCGGAGCGGCGCGCGCTG 60
Oy 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
Db 61 CTGGTCCCGCTCGCGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGG 120
Oy 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGlu 95
Db 121 GTGGTCTAGTTGGGAGCGGCGCATGCTGCCGCCAGAGAGCTCCAGAGGAGGCTG 180
Oy 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluIleSerGlnAsp 115
Db 181 GTGGCGAGAGGAGGACGAGACCCGCTGGAATGATCCCGACAGAGAGAGAGCCAGAT 240
Oy 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
Db 241 CCGGCGCTTCTCTGAGACGACTAGTTGGGCTCGAAGAGAGTGCACCTTAAGCGCGGAA 300

```

```

Oy 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155
Db 301 ACACGGGCTCGAAGACCGATCGCAGCCATTATGAAGTTTCATCCAGACGTCGAGGAC 360
Oy 156 GlyAlaGluAlaGlyValAspGlyThrValSerGlyThrGluGluAlaArgIleAsnSer 175
Db 361 GAGGCCAGGACAGGTGTGACAGGACAGTACGTAGTGGCTGGGAGAGACCAATCAACAGC 420
Oy 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
Db 421 TCAGCGCTCTGGCTGACACACCGCAGATCGGGAGTTTATAGTCACCGGGCTGGGCTG 480
Oy 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215
Db 481 TACTACCTGTACTGTAGGTGACCTTGTAGAGGAGGAGGCTGCTACTGAGAGCTGGAC 540
Oy 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
Db 541 TTGCTGTGGATGTGTGTGCTGGCCCTGCGCTCGGAGAAATTCACAGCCACTGCGCGG 600
Oy 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255
Db 601 ACTTCCCTCGGGGCCACAGCTCCGCTGCGCAGAGTGTGTGGCTGTGGCCCTGCGGCA 660
Oy 256 GlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275
Db 661 GGGTCTCTCCCTGGGATCGGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCTC 720
Oy 276 ThrTyrPheGlyLeuPheGlnValHis 284
Db 721 ACTTACTCGGACTCTTCCAGGTTTCAC 747

RESULT 6
AX180714 898 bp DNA linear PAT 06-AUG-2001
LOCUS AX180714 Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 898)
AUTHORS Wiley,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES
source 1.898
location/Qualifiers
CDS
/note="human TWEAK fusion protein construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/translation="MATSRTSLLAFLGLCLPMLQEGSASSDRMKQIEDKIEILS
KTYHENEIARIKLIGERTRSLSRASLSHOEPAQELVAEEDQDSLELPTEES
ODPAFNLVLRPRSAPGRKTRARAIATAHYEVRPGODGAGAGVGTSGMERA
RINSSPLRYNQICEFTYTRAGIYKIQVHFDSKANYLKLIDLINGVALRCLLE
FSATASSLGPOLKCOVSGIALAPRGSSLRIRLIPMALAKAPLTYFGLEQVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Alignment Scores:
Pred. No.: 1.19e-47 Length: 898
Score: 1066.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.82% Indels: 0
Gaps: 0
DB: 6

```

US-09-245-198a-4 (1-284) x AX180714 (1-898)

QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluValAla 97  
 |||||||  
 Db 250 AGTTGGGGACCCGGGCTGCTGCTGCCCGACAGAGCTGCCACAGAGAGCTGGTGGCA 309

QY 98 GluGluAspGlnAspProSerGluLeuAsnProGlnTrpGluGluSerGlnAspProAla 117  
 |||||||  
 Db 310 GAGAGAGCAGAGACCCGTCGGAAGTGAATCCAGACAGAAAGAACAGAGATCTCGG 369

QY 118 ProHeLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThrArg 137  
 |||||||  
 Db 370 CCTTCCTGAACGACGACTAGTTCGCTCGCAGAGATGACCTAAAGCCGGAACACAGG 429

QY 138 AlaArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGluGlnAspGlyAla 157  
 |||||||  
 Db 430 GCTCGAAGAGCGGATCGCGCCCATTTATGAATTCATCCACAGACCTGGACAGAGAGCG 489

QY 158 GlnAlaGlyValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSerSer 177  
 |||||||  
 Db 490 CAGCAGAGTGTGAGCAGGAGCAGTGTGTGGAGAGAACAGACAGATCAGAGCTCCAGC 549

QY 178 ProLeuArgTyrAsnArgGlnIleGlyGluPheIleValIleThrArgAlaGlyLeuTyrTyr 197  
 |||||||  
 Db 550 CCTTCGCTGACAAACCGCCAGATCGGGAGATTATAGTACCCCGGCTGCTACTAC 609

QY 198 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 217  
 |||||||  
 Db 610 CTGACTGTGAGTGCACTTTGATGAGGAGGAGGAGGAGTGTCTACCTGAGACTGTG 669

QY 218 ValAspGlyValIleAlaLeuValArgCysLeuGluGluPheSerAlaThrAlaAlaSerSer 237  
 |||||||  
 Db 670 GTGATGTGTGTGCTGGCTGCTGCTGCTGAGGAAATCTCAGACACTGCGGCCAGTTC 729

QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuAlaLeuArgProGlySer 257  
 |||||||  
 Db 730 CTCGGGCCCCAGCTCCGCTGTGCGAGGTGTGTGGCTGTGGCCCTCGGCGAGGCTC 789

QY 258 SerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyr 277  
 |||||||  
 Db 790 TCCCTGGGATCCGACCCCTCCCTGGGCGCATCTCAAGAGGTGCCCTTCTCCTAC 849

QY 278 PheGlyLeuPheGlnValHis 284  
 |||||||  
 Db 850 TTCGAGCTCTCCAGGTTCAC 870

RESULT 7  
 AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS Mus musculus TWEAK mRNA, partial cds.  
 DEFINITION AF030100  
 ACCESSION AF030100.1 GI:2707220  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1168)  
 Chicheportiche Y., Bourdon P.R., Xu H., Han Y.M., Scott H.,  
 Hession C., Garcia I. and Browning J.L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)  
 J98070415  
 9405449  
 2 (bases 1 to 1168)  
 Chicheportiche Y., Bixler S., Tizard R. and Browning J.  
 Direct Submission  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
 Cambridge, MA 02142, USA  
 FEATURES  
 Location/Qualifiers  
 1..1168

CDS

/organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_type="peritoneal macrophages"  
 <1..679  
 /note="ligand in the TNF family; secreted protein"

BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,79e-45 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 DB: 10 Gaps: 0

US-09-245-198a-4 (1-284) x AF030100 (1-1168)

QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValIleSerLeuGly 80  
 |||||||  
 Db 5 CTGAGCCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 64

QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluValAlaGluGluAsp 100  
 |||||||  
 Db 65 AGCTGGGCAAGCAGCTGTGCCAGAGAGCTCTCAGAGAGAGCTGACAGAGAGAGCC 124

QY 101 GlnAspProSerGluLeuAsnProGlnTrpGluGluSerGlnAspProAlaProPheLeu 120  
 |||||||  
 Db 125 CGGAGGCCCTGAACTGAATCCAGACAGAGAAACCCAGAGTGTACTTCTTG 184

QY 121 AsnArgLeuValArgProArgArgSerAlaProGlyLysAlaArgGlyThrArgAlaArg 140  
 |||||||  
 Db 185 GACACACTAGTCCGCGCTCGAAGAGTCTCTTAAGGCCGAGAGCCGCTCGCGCA 244

QY 141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
 |||||||  
 Db 245 GCTATTGAGCCCATTTATGAGGTTCATCTCGGCGAGAGAGAGTGCACAAAGAGGT 304

QY 161 ValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSerProLeuArg 180  
 |||||||  
 Db 305 GTGATGGGACAGTAGTGTGGGAGAGACCAAAATCAACAGCTCCAGCCCTGCGC 364

QY 181 TyrAsnArgGlnIleGlyGluPheIleValIleThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
 |||||||  
 Db 365 TACGACCCGACAGATTGGGGAATTAAGTCAATCAGAGCTGGGCTCTACTACTGACT 424

QY 201 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
 |||||||  
 Db 425 CAGGCGCACTTGTATGAGGAGAAAGGCTGTACTGAAAGCTGCTGTTGAAGCGT 484

QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerLeuGlyPro 240  
 |||||||  
 Db 485 GTGCGGCGCTGGCGCTGGCGTGAAGATTTCCAGCCACAGCAAGAGTCTCTGCGCC 544

QY 241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260  
 |||||||  
 Db 545 CAGCTCCGTTGTGCGCAGAGTGTGCGGTGTGGCGCTGGCGGCGCAGAGGTCTCCCTTCG 604

QY 261 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280  
 |||||||  
 Db 605 ATCCGACACCTCCCTGGGCTGACATTTAAGCTGCGCCCTTCTTAACCTACTTGGACTC 664

QY 281 PheGlnValHis 284  
 |||||||



Db 665 TTTCAGTTCAC 676

RESULT 8  
AC016876/c 177703 bp DNA linear HTG 06-AUG-2002  
LOCUS Homo sapiens clone RP11-186B7, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
DEFINITION unordered pieces.  
ACCESSION AC016876  
VERSION AC016876.5 GI:21313830  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_PULLTOP; HTGS\_ACTIVEPIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 177703)  
Birren,B., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome, clone RP11-186B7  
Unpublished  
2 (bases 1 to 177703)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Bouhgaalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., Dearlano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,  
Gallagan,J., Gardyna,S., Grant,G., Hages,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Margulis,N.,  
McBreen,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Mortow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 177703)  
Birren,B., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Barns,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouhgaalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Gallagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,C., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Milova,T., Mienga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Plunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Teste,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zemke,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 2, 2002 this sequence version replaced gi:15421989.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3849  
Center clone name: 186\_B\_7  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces

is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 48645: contig of 48645 bp in length  
\* 48646 48745: gap of 100 bp  
\* 48746 58277: contig of 9532 bp in length  
\* 58278 58377: gap of 100 bp  
\* 58378 65804: contig of 7427 bp in length  
\* 65805 65904: gap of 100 bp  
\* 65905 79793: contig of 13889 bp in length  
\* 79794 79893: gap of 100 bp  
\* 79894 99493: contig of 19600 bp in length  
\* 99494 99593: gap of 100 bp  
\* 99594 111049: contig of 11456 bp in length  
\* 111050 111149: gap of 100 bp  
\* 111150 125020: contig of 13871 bp in length  
\* 125021 125120: gap of 100 bp  
\* 125121 145109: contig of 19989 bp in length  
\* 145110 145209: gap of 100 bp  
\* 145210 169458: contig of 24249 bp in length  
\* 169459 169558: gap of 100 bp  
\* 169559 177703: contig of 8145 bp in length.  
Location/Qualifiers  
1. 177703  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-186B7"  
/clone\_11b="RP11-11 Human Male BAC"  
BASE COUNT 42862 a 45967 c 46568 g 41254 t 1052 others  
ORIGIN  
Alignment Scores:  
Pred. NO.: 1.39e-20 Length: 177703  
Score: 589.50 Matches: 125  
Percent Similarity: 65.62% Conservative: 1  
Best local Similarity: 65.10% Mismatches: 0  
Query Match: 40.82% Indels: 66  
DB: 2 Gaps: 1  
US-09-245-198a-4 (1-284) x AC016876 (1-177703)  
QY 159 AlaGlyValaSpGlyThyValSerGlyTyrPouLunAlaArgIleAsnSerSerPro 178  
Db 37636 TCAGGTGTGGACGGGACAGTGTGGTGGGAGGAGGACCAATCACAGCTCCAGCCCT 37577  
QY 179 LeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeu 198  
Db 37576 CTGGCGTACACACCGCCAGATGGGGAGATTATAGTACCCCGGCTGGCTTACTACTCCTG 37517  
QY 199 TyrCys-----  
Db 37516 TACGTGTAGTAAAGCCCATCTGGCTGCATGGTAAAGCACTAGAGAGAGGAGGCT 37457  
QY 200 -----  
Db 37456 TTGCAGAGAGAGTGGGGGACAGCTAGAGGGCTGGAGGGTGGGCTTGGGTGGG 37397  
QY 200 -----  
Db 37396 ATGGAGATCCCTGCTGCTGCTGAGGAAATTGAAATTGAGGCGAGGCGAGGAGGCTGG 37337  
QY 201 -----GlnValHisPheAspGluGlyLysAlaValTyrLeu 213  
Db 37336 ACTGGGCTGTGTGTCACACCCAGGTGACATTTGATGAGGAGGAGGAGGCTGCTACCGAA 37277  
QY 213 sLeuAspLeuValaSpGlyValIleValaLeuArgCysLeuGluIupheserAlaIarh 233  
Db 37276 GCTGAGT 37217  
QY 233 rAlaAlaserIerleuGlyProGlnleuArgLeuGlyCysGlnValSerGlyLeuLeuAlaIe 253

```

|||||
Db 37216 TGGGGAGATTCCCTGGGGCCAGCTCCGCCCTCCAGTGTCTGTGGCCCT 37157
Oy 253 uatgpgoglyserSerleuArgrlleArgrhleuprotRpalahleuylsAlaIapR 273
Db 37156 GCGGCCAGGGTCTCTCCCTCGGATCCGACCTCCCTGGCCCACTCAAGGTCGCC 37097
Oy 273 ophleuthrTyrrPheglyLeupheglnValhIs 284
Db 37096 CTTCCTCACCCTACTCGACTCTTCAGGTTCAC 37063

RESULT 9
AC127470 215795 bp DNA linear HMG 17-JUL-2002
LOCUS Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15
DEFINITION
AC127470
AC127470.1 GI:21886866
KEYWORDS HMG, HMG_PHASE1, HMG_DRAFT.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1 (bases 1 to 215795)
REFERENCE
AUTHORS Akhter, N., Antonellis, A., Ayala, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
Lee-Ilh, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Marquiles, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,
McConkey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Startipop, S., Thomas, J.W.,
Thomas, P.D., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 215795)
Green, E.D.
Direct Submission
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomt Circle, Gaithersburg, MD 20877, USA
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: cms
Center clone name: 145D13
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-contrigs
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-contrigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 3593: contig of 3593 bp in length
* 3594 3593: gap of unknown length
* 8001 8001: contig of 4308 bp in length
* 8002 8101: gap of unknown length

```

```

* 8102 15639: contig of 7538 bp in length
* 15640 15739: gap of unknown length
* 15740 24775: contig of 9036 bp in length
* 24776 24876: gap of unknown length
* 24876 30150: contig of 5275 bp in length
* 30151 30250: gap of unknown length
* 30251 39312: contig of 8962 bp in length
* 39313 39312: gap of unknown length
* 39313 49342: contig of 10030 bp in length
* 49343 49443: gap of unknown length
* 49443 60956: contig of 11514 bp in length
* 60957 61056: gap of unknown length
* 61057 74520: contig of 13464 bp in length
* 74521 74620: gap of unknown length
* 74621 89522: contig of 14902 bp in length
* 89523 89622: gap of unknown length
* 89623 105537: contig of 15515 bp in length
* 105538 105637: gap of unknown length
* 105638 121197: contig of 15560 bp in length
* 121198 121297: gap of unknown length
* 121298 138883: contig of 17586 bp in length
* 138884 138983: gap of unknown length
* 138984 166802: contig of 27819 bp in length
* 166803 166902: gap of unknown length
* 166903 215795: contig of 48893 bp in length.
Location/Qualifiers
1. 215795
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-145D13"
/clone="lib-RP43"
1. 3593
/note="assembly-fragment"
3594..8001
/note="assembly-fragment"
8102..15639
/note="assembly-fragment"
15740..24775
/note="assembly-fragment"
24876..30150
/note="assembly-fragment"
30251..39212
/note="assembly-fragment"
39313..49342
/note="assembly-fragment"
49443..60956
/note="assembly-fragment"
61057..74520
/note="assembly-fragment"
74621..89522
/note="assembly-fragment"
89623..105537
/note="assembly-fragment"
105638..121197
/note="assembly-fragment"
121298..138883
/note="assembly-fragment"
138984..166802
/note="assembly-fragment"
166903..215795
/note="assembly-fragment"

BASE COUNT 52160 a 54859 c 53752 g 53513 t 1511 others
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-20 Length: 215795
Score: 589.50 Matches: 125
Percent Similarity: 65.62% Conservative: 1
Best Local Similarity: 65.10% Mismatches: 0
Query Match: 40.82% Indels: 66
DB: 2 Gaps: 1
US-09-245-198a-4 (1-284) x AC127470 (1-215795)

```



clone\_end:sp6  
vector\_side:right"  
BASE COUNT 41026 a 46997 c 46681 g 41924 t 927 others  
ORIGIN

Alignment Scores:  
Pred. No.: 6.8e-19 Length: 177555  
Score: 557.00 Matches: 130  
Percent Similarity: 52.82% Conservative: 20  
Best Local Similarity: 45.77% Mismatches: 45  
Query Match: 38.57% Indels: 89  
DB: 2 Gaps: 4

US-09-245-198a-4 (1-284) x AC130192 (1-177555)

67 LeuAlaCysLeuGlyLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86  
DB 114760 TTGCTCTGTAATCTTTCTGTCATGATGATGATTTTTCACCTCATTCTATCACA 114701  
87 AlaGlnGluPro-----AlaGlnGluGluLeuValAlaGluGluAspGlnAspPro 103  
DB 114700 CTTTGGGCTCCAGAGGCGGTGCTCCCTTCCTTGCCTAGCCATCTGTATCATCCG 114641  
104 SerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 123  
DB 114640 GGTACGGTGGGCGCCAGGTTGATCAATTCAGTGGCCCGGGGATTTATTA----- 114590  
124 ValArgProArgArgSerAlaProGlyArgLysArgLysThrArgAlaArgAlaIleAla 143  
DB 114589 -----GAAAGCTCTGTAAAGGACTTTC----- 114566  
144 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163  
DB 114565 -----CCTGTCTGTACATCATCCACTCATCGTGTAAACGGG 114530  
164 ThrValSerLysTyrPglGluGluAlaArgLysSerSerSerProLeuArgTyrAsnArg 183  
DB 114529 ACGGTGATGTGGTGGGAGGAGGCAAAATCAACAGCTCCAAACCATGCGTATGACCGC 114470  
184 GlnIleGlyLysPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys----- 200  
DB 114469 CAAAGGGGCAATTTTGGTGCACCGGGGAGGCTGTACTACTGTCTCAGTAAAGC 114410  
200 ----- 200  
DB 114409 CCCAGCTGCTCCCGGCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 114350  
200 ----- 200  
DB 114349 GGGCCCAAGTTAAAGTGGGGAAGGGGAGGTTGTGCTGGAAGGATGCGCGGCTCTT 114290  
201 -----G1 201  
DB 114289 GGGAAACCTGAGACAGGGGTCAAGGGCAAGAGGCTGAGCTCCGCCCCCCCCCA 114230  
201 nValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyVal 221  
DB 114229 GGTGCACTTTGATGAGGGGAGGCGGCTGTACTGAAAGCTGCTGTGTGAACGACAC 114170  
221 LeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyProG1 241  
DB 114169 GGTGGGCTGCTGCTGCTGAGGAGGATTTCTGCGACGGCGGAGTCCCTCGGGCCCA 114110  
241 nLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg1 261  
DB 114109 GCTCCGCTCTGCGCAGGTGTCTGGGCTATTGCCCCCTGCGGCGGCTCTCCCTGCGGAT 114050  
261 eArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPh 281  
DB 114049 CCGCACCTCTCCCTGGGCGCACCTCAAGGCTGCGCCCTTCTCACCGTCTTGGAGCTCT 113990  
281 eGlnValHis 284  
|||||

DB 113989 CCAAGTTCAC 113980

RESULT 11  
LOCUS AC126925/c  
DEFINITION Canis familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 15  
unordered pieces.  
AC126925  
AC126925.1 GI:21724102  
VERSION  
KEYWORDS HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
1 (bases 1 to 161428)

# REFERENCE

ADTHORS  
Akhter N., Antonellis A., Ayala K., Beckstrom-Sternberg S.M.,  
Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C.,  
Brooks S., Dietrich N.L., Granle S., Guan X., Gupta J.,  
Haghighi P., Hansen N., Ho S.-L., Idol J.R., Karlins E., Latic P.,  
Lee-Hin S.-O., Legaspi R., Maduro O.L., Maduro V.B.,  
Margulies E.H., Mastello C., Maskell B., Mastrian S.D.,  
McCluskey J.C., McDowell J., Paguirigan C., Pearson R.,  
Portnoy M.E., Prasad A., Schueler M.G., Stantiridop S., Thomas J.W.,  
Thomas P.J., Touchman J.W., Tsurgon C., Vogt J.L., Walker M.A.,  
Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 161428)

# TITLE

# REFERENCE

# ADTHORS

# TITLE

# JOURNAL

# COMMENT

# Center: NIH Intramural Sequencing Center

# Center code: NISC

# Web site: http://www.nisc.nih.gov

# Contact: nisc-zoo@ngt.nih.gov

# Project Information

# Center project name: 332E11

# Center clone name: 332E11

# Summary Statistics

# Sequencing vector: plasmid: n/a; 100% of reads

# Chemistry: Dye-terminator Big Dye; 100% of reads

# Assembly program: Phrap; version 0.990319

# Consensus quality: 148712 bases at least Q40

# Consensus quality: 151471 bases at least Q30

# Consensus quality: 153125 bases at least Q20

# Insert size: 15200; agarose-ftp

# Insert size: 160028; sum-of-contigs

# Quality coverage: 5.36x in Q20 bases; sum-of-contigs

# Quality coverage: 5.10x in Q20 bases; sum-of-contigs

# NOTE: This is a 'working draft' sequence. It currently

# consists of 15 contigs. The true order of the pieces

# is not known and their order in this sequence record is

# arbitrary. Gaps between the contigs are represented as

# 'runs of N', but the exact sizes of the gaps are unknown.

# This record will be updated with the finished sequence

# as soon as it is available and the accession number will

# be preserved.

# 1

# 4735

# 4835

# 7325

# 7426

# 12262

# 12362

# 19032

# 19132

# 27850

# 27950

# 31802

# 31901

# 40781

# contig of 4734 bp in length

# gap of unknown length

# contig of 2491 bp in length

# gap of unknown length

# contig of 4836 bp in length

# gap of unknown length

# contig of 6670 bp in length

# gap of unknown length

# contig of 8718 bp in length

# gap of unknown length

# contig of 3852 bp in length

# gap of unknown length

# contig of 8880 bp in length

```

/organism="Canis familiaris"
/db_xref="taxon:9615"
/clone="RP81-332E11"
/clone_id="RP81"

1..4734
misc_feature
    /note="assembly-fragment"
    4835..7325
misc_feature
    /note="assembly-fragment"
    7425..12261
misc_feature
    /note="assembly-fragment"
    12362..19031
misc_feature
    /note="assembly-fragment"
    19132..27849
misc_feature
    /note="assembly-fragment"
    27950..31801
misc_feature
    /note="assembly-fragment"
    31902..40781
misc_feature
    /note="assembly-fragment"
    40882..49783
misc_feature
    /note="assembly-fragment"
    49884..59150
misc_feature
    /note="assembly-fragment"
    59251..68679
misc_feature
    /note="assembly-fragment"
    68780..78384
misc_feature
    /note="assembly-fragment"
    79085..93446
misc_feature
    /note="assembly-fragment"
    93547..102175
misc_feature
    /note="assembly-fragment
    clone_end:SP6
vector_slide:right"
102176..131685
misc_feature
    /note="assembly-fragment"
    131786..161428
BASE COUNT      39586 a 41863 c 41333 g 37340 t      1406 others
ORIGIN

Alignment Scores:
Pred. No.:      7,39e-19      length:      161428
Score:          555.50      Matches:      118
Percent Similarity: 64.92%      Conservative: 6
Best Local Similarity: 61.78%      Mismatches: 9
Query Match:      38.47%      Indels:      58
DB:              2          Gaps:      1

```

Db	Accession	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	COMMENT
Db 42746	AAATCAACAGACTCCAAACCCACTGGCTATGACCGCCAGAGGGGGAATTATATCAAG	42746	AAATCAACAGACTCCAAACCCACTGGCTATGACCGCCAGAGGGGGAATTATATCAAG	42687						
Db 42686	CGGCTGGGCTCTACTACTCTACTGACGAGTAAGCCCACTGGCTCCACGGTAGGCA	42686	CGGCTGGGCTCTACTACTCTACTGACGAGTAAGCCCACTGGCTCCACGGTAGGCA	42627						
Db 42626	GGAAGCGGGGGCCCGAGGTTAGGAGAGAGAGAGGGTTGGTTTCAGGGGGCAGTTGGG	42626	GGAAGCGGGGGCCCGAGGTTAGGAGAGAGAGAGGGTTGGTTTCAGGGGGCAGTTGGG	42567						
Db 42566	GTGGGGAGACCTTAGGGTCGGGGGCCCGAGTGGATCCTGGGGGTACCAAGACCCCGACTCT	42566	GTGGGGAGACCTTAGGGTCGGGGGCCCGAGTGGATCCTGGGGGTACCAAGACCCCGACTCT	42507						
Db 42506	GGCGCTCCCTCCGCGGCCCGCCAGGTGCACCTTGATGAGGGGAAGCGCTGCTACCTGAAGCT	42506	GGCGCTCCCTCCGCGGCCCGCCAGGTGCACCTTGATGAGGGGAAGCGCTGCTACCTGAAGCT	42447						
Db 42446	GACCTTGTGTGTGATGACGCGCCCTGGCCCTCGCGCTCGTGAAGATTCCTCGCACAGC	42446	GACCTTGTGTGTGATGACGCGCCCTGGCCCTCGCGCTCGTGAAGATTCCTCGCACAGC	42387						
Db 42386	CGCCACGACGCTGGGGCCCCCAGCTCCGCTCTGCCAATGCTCTGGGCTGTGGCCCTCCG	42386	CGCCACGACGCTGGGGCCCCCAGCTCCGCTCTGCCAATGCTCTGGGCTGTGGCCCTCCG	42327						
Db 42326	GCCCGGGGCTCCCTCGCGGATCCGACCCCTCCCTCCGCGCCATCTCAAGGCGCCCTT	42326	GCCCGGGGCTCCCTCGCGGATCCGACCCCTCCCTCCGCGCCATCTCAAGGCGCCCTT	42267						
Db 42266	CCTTACTACTCTCGACCTCTCCAGCTTAC	42266	CCTTACTACTCTCGACCTCTCCAGCTTAC	42236						
Db 42236	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42206	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42186	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42166	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42146	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42126	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42106	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42086	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42066	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42046	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42026	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 42006	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41986	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41966	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41946	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41926	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41906	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41886	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41866	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41846	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41826	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41806	AC069459	AC069459	AC069459	203083	bp	DNA	linear	HTG	27-JUN-2001	
Db 41786	AC069459	AC069459	AC069459	203083	bp	DNA</				

```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: MAFU
Center clone name: RP23-168P5
-----
Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-primer Bodily: 48% of reads
Dye-terminator Big Dye: 5% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210556; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 62152: contig of 62152 bp in length
* 62153 62252: gap of unknown length
* 62253 118772: contig of 56520 bp in length
* 118773 118872: gap of unknown length
* 118873 148824: contig of 30052 bp in length
* 148925 149024: gap of unknown length
* 149025 167231: contig of 18207 bp in length
* 167232 167331: gap of unknown length
* 167332 189907: contig of 22576 bp in length
* 189908 190007: gap of unknown length
* 190008 196537: contig of 6530 bp in length
* 196538 196637: gap of unknown length
* 196638 203083: contig of 6446 bp in length.
*
Location/Qualifiers
source 1..203083
          /organism="Mus musculus"
          /db_xref="taxon:10090"
          /chromosome="11"
          /clone="RP23-168P5"
BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others
ORIGIN
Alignment Scores:
Pred. No.: 9,31e-19 Length: 203083
Score: 555.50 Matches: 117
Percent Similarity: 67.22% Conservative: 4
Post Local Similarity: 65.00% Mismatches: 5
Query Match: 38.47% Indels: 54
DB: 2 Gaps: 1
US-09-245-198A-4 (1-284) x AC069459 (1-203083)
OY 159 AlAGlyyAlAGcylThrAlASerGlyTTPglUGluAlArGlLeAsnSerSerPro 178
Db 43305 TCAGGTGTGGATGGACACTGTGCTTGGAAGAACAACCAATCAACAGCTCCACCCTC 43246
OY 179 LeuArgTyrrAsnArgGlnIleGlyGUphellEValTTThrArAlaGlyLeuTyrrIeu 198
Db 43245 CTGCCCTACGACCGCACAGATTGGGGAAATTACAGTCATCATGAGGCTGGGCTACTACTCTG 43186
OY 199 TyrrCyS----- 200
Db 43185 TACTGTCAAGGTAAAGCCCTGCCTGTTCAGGGGACAGCAAGAGCTTAAGGGGAGGAGGA 43126
OY 200 ----- 200

```

	Db	43125	CTGGCAGGAATTTGGTGGGAGGAGACACTTCGCACATCAAGAAGAAATGCCCTTGCTT	43066	
Oy	201	-----	glnvalhisphaas	205	
Db	43065	TCCGATGAGGAGCGCAGAGGTTCGATTTCCTTCCTCTCTCTGTCACAGGTGCACACTTCCA	43006		
Oy	205	pglugllylsalaVal1TYrLeuLysLeuAspLeuValaspGlYValAlaLeuAr	225		
Db	43005	TGAGGGAAGGCTGTCTACCTGAACTGACTGTGTGTGAACGGTGTCTGGCCCTCGC	42946		
Oy	225	gcysleuGlugluPheserAlarThrAlaIAsSerSerLeuGlYProGlnLeuAylGlcY	245		
Db	42945	CTGCCTGGAAAGATTTCTCAGCGCACAGCAAAGCTCTCTCGGGGCCCAAGCTCCGTTGTG	42886		
Oy	245	sglnValserGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThyleuPr	265		
Db	42885	CCAGGTGTCTGGGCTGGTGGCCGCTGGCCAGGGCTTCCTTCGGAATCCGACACTTCCC	42826		
Oy	265	oTrpalHisLeuLysAlaAlaProPheLeuThyrPheGlyLeuPheGlnValHis	284		
Db	42825	CTGGGCTCATCTTAAGGCTGCCCTTCTCTAACCTACTTTGGACCTTTCAAGTTAC	42768		
RESULT 13					
AL603707/c					
LOCUS		234182 bp	DNA	linear ROD 17-NOV-2001	
DEFINITION		Mouse DNA sequence from clone RP23-422L16 on chromosome 11,			
ACCESSION		complete sequence.			
VERSION		AL603707			
KEYWORDS		AL603707.5 GI:17017790			
SOURCE		HTG.			
ORGANISM		house mouse.			
		Mus musculus			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE		1 (bases 1 to 234182).			
JOURNAL		Pearce/A.			
COMMENT		Direct Submission Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonesrequest@sanger.ac.uk On Nov 20, 2001 This sequence version replaced gi:16603765. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORKPEP; Information on the WORKPEP database can be found at <a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a> RP23-422L16 is from the RPIc-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <a href="http://www.chori.org/Bacpac/home.htm">http://www.chori.org/Bacpac/home.htm</a> VECTOR: pBACE3.6 This sequence is the entire insert of clone RP23-422L16. Location/Qualifiers 1..234182 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="11" /clone="RP23-422L16" /clone_1id="RPIc-23" complement(84050...84131)			

```

/Note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
BASE COUNT      59310 a 56824 c 57519 g 60529 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	1.08e-18
Score:	555.50
Percent Similarity:	67.22%
Best local Similarity:	65.00%
Query Match:	38.47%
DB:	10
Length:	23418
Matches:	117
Conservative:	4
Mismatches:	5
Indels:	54
Gaps:	1

US-09-245-198A-4 (1-284) x AL603707 (1-234182)

QY	159	179	198
QY	AlaGlyValAlaPglYThhValSerCylYrpgLuglualAargILAsnSerSerPro	178	
Db	66609 TCAGGTGCGATGGGACAGTGTGCTGGGAGAGACCMAATCAACGCTCCAGCCCT	66550	
QY	LeuAArgYrYAsnArgGlnILeolYglubheILeValIThrArgAlaGlyLeuYrYrIleu	198	
Db	66549 CTGGCCTCAGCAACCGGCAATTTGGGGAATTTACAGTCATCAGGGCTGGGCTACTACCTG	66430	
QY	199 TyrCys	200	
Db	66489 TACTGTCAAGTAAGCCCTGCTGCTTCAGGGGACAGACCAAGGCTAAGGGAGGAGGA	66430	
QY	200	200	

Db 66429 CTGGCAAGGAATGGGTGGGAGAGAGGAGACTCTGCGACTCATGAAGGAAGATGCCCTGGCTT 66370

201 -----GlnValHispheas 205

UU 00303 1CGAGGGAGGCGAGGG11GGATT1GCT1GCC11CT1CTGT1CCAGG1GCAC1TTGA 00310

203 potuolypysniavaiiyneulysneuspreuleuvalasporiyalileualaleuar 223

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

CC BY-NC-ND 4.0 International license.

[illegible]

## RESULT 1A

LOCUS	AC126921	153553 bp	DNA	linear	HTG 10-TM-2
AC120921					

DEFINITION  
BOS CAUDUS CIONE RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered  
pieces.

REVISION	AC126921
VERSION	AC126921.1
	GI:21724098

SOURCE	COW.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

Bovidae; Bovinae; Bos.

**AUTHORS** Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,

Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,

Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,

McCloskey, J. C., McDowell, J., Paguirigan, C., Pearson, R.,  
Postman, M. D.

Thomas, P. J., Touchman, J. W., Tsurean, C., Vogt, J. L., Walker, M. A.,  
 Wetherby, K. D., Wiggins, L., Young, A., Zhang, L. H. and Green, E. D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 153553)  
 Green, E. D.  
 Direct Substition  
 Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 ----- Genome Center

```

Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@harl.nih.gov
-----
Project Information
-----
Center project name: ddi
Center clone name: 045D24
-----
Summary Statistics
-----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 14606 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-configs
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.77x in Q20 bases; sum-of-configs

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 conigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the conigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	1	2313:	contig of 2312 bp in length
*	2313	2413:	gap of unknown length
*	2413	5841:	contig of 3429 bp in length
*	5842	5841:	gap of unknown length
*	5842	8435:	contig of 2494 bp in length
*	8436	8535:	gap of unknown length
*	8536	15799:	contig of 7264 bp in length
*	15800	15889:	gap of unknown length
*	15800	25224:	contig of 9325 bp in length
*	25225	32504:	gap of unknown length
*	32525	32504:	contig of 7180 bp in length
*	32805	32604:	gap of unknown length
*	32805	40970:	contig of 8366 bp in length
*	40971	41070:	gap of unknown length
*	41071	56590:	contig of 15520 bp in length
*	56591	56590:	gap of unknown length
*	56591	73763:	contig of 17079 bp in length
*	73770	73869:	gap of unknown length
*	73870	90855:	contig of 16590 bp in length
*	90860	90955:	gap of unknown length
*	90960	111428:	gap of unknown length
*	111429	111528:	gap of unknown length
*	111529	153535:	contig of 42025 bp in length

FEATURES	Location/Qualifiers
source	1. .153553

```
/organism="Bos taurus"
/db_xref="taxon:9913"
```

```

/clone="RP42-45D24"
/clone_1lb="RP42"

```

```
misc_feature 1. .2312
              /note="assembly_fragment"
              2412 2041
```

```
misc_feature      2413. :3841
/note="assembly_fragment
of one codon"

```

```
misc feature
5943  8435
      clone_end:1/
      vector_side:left"
```

CC40. 12460  
313337-207W







Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project information

Center project name: GUXG

Center clone name: CH230-320N23

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 112720 bases at least Q40

Consensus quality: 11666 bases at least Q30

Consensus quality: 119165 bases at least Q20

NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      1022: contig of 1022 bp in length
1023      1122: gap of unknown length
1123      2394: contig of 1272 bp in length
2395      2494: gap of unknown length
2495      3663: contig of 1169 bp in length
3664      3763: gap of unknown length
3764      4879: contig of 1116 bp in length
4880      4979: gap of unknown length
4980      6787: contig of 1808 bp in length
6788      6888: gap of unknown length
6888      8192: contig of 1305 bp in length
8193      8292: gap of unknown length
8292      9461: contig of 1169 bp in length
9462      9561: gap of unknown length
9562      11435: contig of 1874 bp in length
11436      11535: gap of unknown length
11536      13535: contig of 2000 bp in length
13536      13635: gap of unknown length
13636      17323: contig of 3688 bp in length
17324      17423: gap of unknown length
17424      20406: contig of 2983 bp in length
20407      20506: gap of unknown length
20507      23445: contig of 2939 bp in length
23446      23545: gap of unknown length
23546      25600: contig of 2055 bp in length
25601      25700: gap of unknown length
25701      28049: contig of 2349 bp in length
28050      28149: gap of unknown length
28150      30474: contig of 2325 bp in length
30475      30574: gap of unknown length
30575      35498: contig of 4924 bp in length
35499      35598: gap of unknown length
35599      39284: contig of 3686 bp in length
39285      39384: gap of unknown length
39385      41870: contig of 2586 bp in length
41871      42070: gap of unknown length
42071      45659: contig of 3589 bp in length
45660      45759: gap of unknown length
45760      50332: contig of 4573 bp in length
50333      50432: gap of unknown length
50433      54695: contig of 4263 bp in length
54696      54795: gap of unknown length
54796      58289: contig of 3494 bp in length
58290      58389: gap of unknown length
58390      63599: contig of 5210 bp in length
63600      63699: gap of unknown length
63700      70305: contig of 6606 bp in length
70306      70405: gap of unknown length
70406      76123: contig of 5718 bp in length
76124      76223: gap of unknown length

```

```

*      76224      84961: contig of 8738 bp in length
*      84962      85061: gap of unknown length
*      85062      93615: contig of 8553 bp in length
*      93615      93715: gap of unknown length
*      93715      103352: contig of 9638 bp in length
*      103352      103452: gap of unknown length
*      103452      110299: contig of 6847 bp in length
*      110299      110399: gap of unknown length
*      110399      116180: contig of 5781 bp in length
*      116180      116280: gap of unknown length
*      116280      125484: contig of 9204 bp in length
*      125484      125585: gap of unknown length
*      125585      138792: contig of 13208 bp in length.

```

FEATURES

source

1. 138792  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-320N23"

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others

ORIGIN

Alignment Scores:

Pred. No.: 1.7e-15

Score: 489.50

Percent Similarity: 64.50%

Best Local Similarity: 60.36%

Query Match: 33.90%

DB: 2

Gaps: 1

US-09-245-198a-4 (1-284) x AC119115 (1-138792)

QY 160 GLYVALASPGLYTHYVALSERGLYTRPGIUGLUALAARGILASNSERSESRPROLEU 179

DB 137345 GGTCGATGGAGACGTGTCGTGGAGAGACCAATATACACGTCCAGCCCTCG 137404

QY 180 ARGTYASNARGGLINLEGLYGLUPHEILEVALTHIRAGLAGLYLEUTYTYR 199

DB 137405 CGCTAGACGCCGACAGTTGGGATTTAGGTCATCAGGCGTGTCTACACCTGAC 137464

QY 200 CYS----- 200

DB 137465 TFCAGGTAAGCCCTGGCTCCATGAGTACAGCATGCTAAGGGAGGAGGCTTGCA 137524

QY 200 ----- 200

DB 137525 AGGATGGTGGAGTGGGAGAACCTGGGTTTCATGACGAGAGCCCTGATTCGATG 137584

QY 201 ----- 201

DB 137585 AGGCGAGGAGAGGTCGTGATTTTCGTGTTCTCTCTCCAGGTGACCTTGATG 137644

QY 207 GLYLYSALAVALTYRLEULYSLEULSPLLEULVALASPGLYVALLEULARGCYS 226

DB 137645 GCGAAGGACGTCTACCTGACGTCGTCGTGAATGTCGTGCTGCCCTGCCCTGC 137704

QY 227 LEUGLUGLUPHSESRALALHRAALASERLEULGLYPROGLINLEULARGYCSGLIN 246

DB 137705 CTGGAAGAAFTCTACGCCACACAGCAGCTCTCTGCGGCCACGCTGCTGTCGAG 137764

QY 247 VALISERGLYLEULVALLEULARGPARGISERLEULARGILEARGTHLEUPROFAP 266

DB 137765 GTCGTGGGCTGTGCTGCTGTCGACAGGCTCTCCCTTCGGAATCCGATACCTTGC 137824

QY 267 ALAHLSLEULYSALAALAPROPHLEU 275

DB 137825 ACTCATATTAAAGCGGCTACCTTCGTT 137851

Search completed: May 8, 2003, 03:45:22

Job time : 2530.92 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:02 ; Search time 36,2672 Seconds

(without alignments)  
1043,455 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHKAPFLTYGFLQVH 284

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

A.GeneSeq\_101002:\*

1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1444	100.0	284	19	AAW47525 Homo sapiens tumor
2	1268	87.8	249	20	AAV09369 Human tumor necro
3	1268	87.8	249	21	AAW07526 Amino acid sequenc
4	1268	87.8	249	21	AAV95338 Human PRO207 antit
5	1268	87.8	249	23	AAU86129 Human PRO207 poly
6	1265	87.6	249	19	AAW29745 TNF related endoth
7	1265	87.6	249	19	AAE00891 Human TREPA (TNF r
8	1066	73.8	273	22	AAU03499 TWEAK extracellular
9	1062	73.5	208	20	AAW93590 Human TNFRL3 protei
10	1020	70.6	225	19	AAW47524 Mus musculus tumor

11	1020	70.6	225	21	AAW07527	Amino acid sequenc
12	968	67.0	211	20	AAW93591	Mouse TNFRL3 protei
13	792	54.8	189	19	AAW29746	TNF related endoth
14	792	54.8	189	22	AAE00892	Human UTAflag TREP
15	761	52.7	146	22	AAE00895	Human TREPA (TNF r
16	716	8.0	325	22	ABB67553	Drosophila melanog
17	116	8.0	409	23	AAU77718	Drosophila melanog
18	108.5	7.5	211	21	AAV58216	Canine mature CD15
19	108.5	7.5	260	21	AAV58215	Canine CD154, Can
20	106.5	7.4	254	16	AAW64190	Human 4-1BB-L poly
21	106.5	7.4	254	18	AAW26657	Human 4-1BB ligand
22	106.5	7.4	254	23	ABB75953	Human cytokine 4-1
23	106	7.3	1428	21	AAV97033	Caspase 8-Interact
24	105.5	7.3	406	23	AAU77717	Drosophila melanog
25	104.5	7.2	779	23	ABB07845	Human MAP kinase p
26	104	7.2	409	23	AAU77716	Drosophila melanog
27	104	7.2	1323	15	AAW55248	N-methyl-D-aspartil
28	102	7.1	256	22	AAW25657	Human protein sequ
29	100.5	7.0	647	17	AAW04327	Rat peltin. Ratu
30	100	6.9	220	22	AAW62340	GP120 V3 loop-CD15
31	99	6.9	574	21	AAV97032	Caspase 8-Interact
32	99	6.9	1008	22	AAW78891	Human protein SEQ
33	99	6.9	1020	22	AAW79875	Human protein SEQ
34	96.5	6.7	1097	22	ABG25655	Novel human diagno
35	96.5	6.7	1631	22	ABG22481	Novel human diagno
36	96	6.6	234	22	AAW62339	Human HIVM-Binding
37	95	6.6	240	23	AAE13680	Human H2A receptor
38	95	6.6	294	13	AAW69956	Human H2A receptor
39	95	6.6	294	19	AAW68292	Human H2A receptor
40	95	6.6	294	22	AAE08737	Human H2A receptor
41	95	6.6	294	22	AAE04425	Human H2A receptor
42	95	6.6	294	22	AAE01992	Human H2A receptor
43	94.5	6.5	1006	22	ABG21178	Novel human diagno
44	94	6.5	876	22	ABG00217	Novel human diagno
45	93.5	6.5	409	22	AAO00076	Human polypeptide

#### ALIGNMENTS

RESULT 1	AAW47525	standard; Protein; 284 AA.
ID	AAW47525	
XX	AAW47525;	
AC	21-JUL-1998	(first entry)
DT		
XX		
DE	Homo sapiens tumor necrosis factor related ligand (TRELL).	
XX		
KW	TRELL; tumor necrosis factor related ligand; tnfr; treatment;	
KW	cancer; autoimmune disease; immune system; stimulation; suppression;	
KW	graft rejection;	
XX		
OS	Homo sapiens.	
XX		
PN	WO9805783-A1.	
PD	12-FEB-1998.	
XX		
PF	07-AUG-1997;	97WO-US13945.
XX		
PR	18-MAR-1997;	97US-0040820.
PR	07-AUG-1996;	96US-0023541.
XX	18-OCT-1996;	96US-0028515.
PA	(BIOJ ) BIOGEN INC.	
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.	
XX		
PI	Browning JL, Chicheportliche Y;	
XX		
DR	WPI, 1998-145619/13.	
DR	N-PSDB; AAV18600.	

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 50-51; 69pp; English.

CC The sequence is that of human tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. It's coding sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.

SQ Sequence 284 AA;

Query Match 100.0%; Score 1444; DB 19; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-129;  
 Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLLDFEISARRLPRISGSDGAVROAPAPMAARRRRGRGEPETALLVPLA 60  
 DB 1 MSLLDFEISARRLPRISGSDGAVROAPAPMAARRRRGRGEPETALLVPLA 60  
 OY 61 LGLGIALACGLLAVSLGSRASLSAQEPAGEELVAEEDDPSELNPOTEESODPAPFL 120  
 DB 61 LGLGIALACGLLAVSLGSRASLSAQEPAGEELVAEEDDPSELNPOTEESODPAPFL 120  
 OY 121 NRVLPRRSAPKGRKTRARRAIAAHYEVHPRPGOGAAGVDGTVSGMEARINSSPLR 180  
 DB 121 NRVLPRRSAPKGRKTRARRAIAAHYEVHPRPGOGAAGVDGTVSGMEARINSSPLR 180  
 OY 181 YNRQIGERTYRAGLYLYCOVHFDEGKAVYIKLDLVDGYALRCLFEFSATASSGLP 240  
 DB 181 YNRQIGERTYRAGLYLYCOVHFDEGKAVYIKLDLVDGYALRCLFEFSATASSGLP 240  
 OY 241 QLRLCOVSGLLALRPGSSLRIRTLPMHLKAPFLTYGFLQVH 284  
 DB 241 QLRLCOVSGLLALRPGSSLRIRTLPMHLKAPFLTYGFLQVH 284

RESULT 2

AA09369  
 ID AA09369 standard; protein: 249 AA.

AC AA09369;

DT 15-JUL-1999 (first entry)

DE Human tumour necrosis factor Apo-3 ligand protein sequence.

KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;

KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

KW cancer.

OS Homo sapiens.

PN WO919490-A1.

PD 22-APR-1999.

PF 09-OCT-1998; 98WO-US21407.  
 PR 17-DEC-1997; 97US-0069862.  
 PR 10-OCT-1997; 97US-0062037.

PA (GETH ) GENENTECH INC.

PI Ashkenazi AJ, Marsters SA, Pittl R;

DR WPI: 1999-287982/24.

DR N-PSDB; AAX56000.

PT New human Apo-3 ligand (a tumour necrosis factor) homologue

PS Claim 1; Fig 1; 74pp; English.

CC The present sequence represents a human tumour necrosis factor (TNF)  
 CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
 CC to induce JNK/SAPK-dependent responses in mammalian cells.

SQ Sequence 249 AA;

Query Match 87.8%; Score 1268; DB 20; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 MAARRSORRRGRGEGPALLVPLALGILACIGLLAVSLGSRASLSAQEPAGEEL 95  
 DB 1 MAARRSORRRGRGEGPALLVPLALGILACIGLLAVSLGSRASLSAQEPAGEEL 95  
 OY 96 VAEEDQDPSSELPQTEESODPAPFLNRLVRRSAPKGRKTRARRAIAAHYEVHPRGOD 155  
 DB 61 VAEEDQDPSSELPQTEESODPAPFLNRLVRRSAPKGRKTRARRAIAAHYEVHPRGOD 120  
 OY 156 GAQAGVDVTGSGMEARINSSPLRYNRQIGEFYTRAGLYLYCOVHFDEGKAVYIKLD 215  
 DB 121 GAQAGVDVTGSGMEARINSSPLRYNRQIGEFYTRAGLYLYCOVHFDEGKAVYIKLD 180  
 OY 216 LLDVGVIALRCLFEFSATASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHLKAPFL 275  
 DB 181 LLDVGVIALRCLFEFSATASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHLKAPFL 240  
 OY 276 TYFGFLQVH 284  
 DB 241 TYFGFLQVH 249

RESULT 3

AAB07526  
 ID AAB07526 standard; protein: 249 AA.

AC AAB07526;

DT 20-OCT-2000 (first entry)

DE Amino acid sequence of a soluble recombinant human TWEAK protein.

KW TWEAK protein; immunological disorder; immune response; inflammation;

KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;

KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.

OS Homo sapiens.

PN WO200042073-A1.

PD 20-JUL-2000.

PF 14-JAN-2000; 2000WO-US01044.

PR 15-JAN-1999; 99US-0116168.

Query Match	Best Local Similarity	100.0%;	Score 1268;	DB 21;	Length 249;
Matches 249;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 36	MAARSSQRRRGRGRRGPTALLVPLALGIGLALACIGLLAVVSLGSRASLSAQEPAGEEL	95			
DB 1	MAARSSQRRRGRGRRGPTALLVPLALGIGLALACIGLLAVVSLGSRASLSAQEPAGEEL	60			
QY 96	VAEEQODSESLNPQEEESODPAFLNRLVPRPRSA PKGKRTARRAIAHYVHPRPGOD	155			
DB 61	VAEEQODSESLNPQEEESODPAFLNRLVPRPRSA PKGKRTARRAIAHYVHPRPGOD	120			
QY 156	GAQAGVDGTVSGWEARINSSPLRNQIGEFITVRAGLYLYLCQVHDEGKAYYLKLD	215			
DB 121	GAQAGVDGTVSGWEARINSSPLRNQIGEFITVRAGLYLYLCQVHDEGKAYYLKLD	180			
QY 216	LIVDGVLLRCLFEESATRAASSLGQRLRCQVSGILLARPGSSLRIRLPMNAHLKAAPFL	275			
DB 181	LIVDGVLLRCLFEESATRAASSLGQRLRCQVSGILLARPGSSLRIRLPMNAHLKAAPFL	240			
QY 276	TYFGGLFOYH 284				
DB 241	TYFGGLFOYH 249				

RESULT 4  
AAAY95338  
ID AAAY95338 standard; Protein: 249 AA.  
AC AAY95338;  
XX 25-SEP-2000 (first entry)  
DT xx  
DE Human PRO207 antitumour protein.  
XX  
XX PRO207; human; antitumour; tumour; therapy; cytosolic;  
XX breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
XX uterine cancer; prostate cancer; lung cancer; bladder cancer;  
XX central nervous system cancer; melanoma; leukaemia; neoplasm.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT Peptide 1..40  
FT /label= Signal\_peptide  
FT Protein 41..249  
FT /label= PRO207

FT	Modified-site	27..33	"N-myristoylation"
FT	/note-	29..35	"N-myristoylation"
FT	Modified-site	36..42	"N-myristoylation"
FT	/note-	45..51	"N-myristoylation"
FT	Modified-site	118..124	"N-myristoylation"
FT	/note-	121..127	"N-myristoylation"
FT	Modified-site	125..131	"N-myristoylation"
FT	/note-	128..134	"N-myristoylation"
FT	Modified-site	139..143	"N-myristoylation"
FT	/note-	10..14	"asn is N-glycosylated"
FT	Modified-site	97..101	"amidation"
FT	/note-	24..35	"amidation"
FT	Peptide	/note-	"prokaryotic membrane lipoprotein lipid"
XX			
PN	WO200037638-A2.		
PD	29-JUN-2000.		
PE	02-DEC-1999;	99MO-US8565.	
XX			
XX	22-DEC-1998;	98US-0113296.	
PR	08-MAR-1999;	99MO-US05028.	
PR	21-APR-1999;	99US-0130232.	
PR	28-APR-1999;	99US-0131445.	
PR	14-MAY-1999;	99US-0134287.	
PR	20-JUL-1999;	99US-0144758.	
PR	26-JUL-1999;	99US-0145698.	
PR	15-SEP-1999;	99MO-US21090.	
PR	15-SEP-1999;	99MO-US21547.	
XX			
PA	(GETH ) GENENTECH INC.		
XX			
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;		
PI	Napier MA, Pitti RM, Wood WI;		
XX			
DR	WPI: 2000-442668/38.		
DR	N-PSSDB; AAAA9717.		
PT			
PT	Newel composition to inhibit neoplastic cell growth or for treating		
PT	tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,		
PT	PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or		
PT	PRO866		
XX	-		
PS	Claim 19; Fig 4; I72pp; English.		
XX			
CC	The present sequence is that of human antitumour protein PRO207,		
CC	as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207		
CC	shows amino acid sequence identity to tumour neurosis factor family		
CC	members, especially human lymphotoxin-beta (23.4%) and human CD40		
CC	ligand (19.8%). Mol. wt. is 27,216. A claimed method for inhibiting		
CC	the growth of a tumour cell comprises exposing the tumor cell		
CC	to PRO179, PRO207, PRO219, PRO221, PRO224, PRO328, PRO301,		
CC	PRO526, PRO362, PRO356, PRO509 or PRO866 (see AA95337-49), their		
CC	agonists or chimeric polypeptides incorporating them. The tumour		
CC	is especially a cancer selected from breast, ovarian, renal,		
CC	colorectal, uterine, prostate, lung, bladder and central nervous		
CC	system cancer, melanoma and leukaemia. Methods for the recombinant		
CC	expression of the antitumour proteins are also provided.		
XX			
SO	Sequence 249 AA;		

Query Match 87.8%; Score 1268; DB 21; Length 249;

Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRSQRRRRGRGEGTALVPLALGCLALACGLLAVVSLGSRASLSAQEPAGEEL 95  
DB 1 MAARRSQRRRRGRGEGTALVPLALGCLALACGLLAVVSLGSRASLSAQEPAGEEL 60  
QY 96 VAEEDDPSELNPTQESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 155  
DB 61 VAEEDDPSELNPTQESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 120  
QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGFEITYRAGLYLYLCQVHDEKAVYLLKD 215  
DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGFEITYRAGLYLYLCQVHDEKAVYLLKD 180  
QY 216 LTVDGVALRCLEESFATRAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHKLKAAAPFL 275  
DB 181 LTVDGVALRCLEESFATRAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHKLKAAAPFL 240  
QY 276 TYFGLEFQVH 284  
DB 241 TYFGLEFQVH 249

# RESULT 5

AAU86129 standard; Protein: 249 AA.

AC AAU86129;  
DT 15-JUL-2002 (first entry)  
DE Human PRO207 polypeptide.  
KW Human: PRO: benign tumour; malignant tumour; lymphoid malignancy;  
KW leukaemia; neuronal disorder; stromal disorder; blastocoeil disorder;  
KW inflammatory disorder; immune disorder; angiosenic disorder;  
KW cytoskeletal; neuroprotective.  
OS Homo sapiens.  
PN WO200153486-A1.  
PD 26-JUL-2001.

XX 11-FEB-2000; 2000WO-US03565.  
XX 08-MAR-1999; 99WO-US05028.  
XX 11-MAR-1999; 99US-123972P.  
XX 11-MAY-1999; 99US-133459P.  
XX 02-JUN-1999; 99WO-US12252.  
XX 22-JUN-1999; 99US-140650P.  
XX 22-JUN-1999; 99US-140653P.  
XX 20-JUL-1999; 99US-144758P.  
XX 26-JUL-1999; 99US-145698P.  
XX 28-JUL-1999; 99US-146222P.  
XX 17-AUG-1999; 99US-149395P.  
XX 31-AUG-1999; 99US-151689P.  
XX 01-SEP-1999; 99WO-US20111.  
XX 15-SEP-1999; 99WO-US21090.  
XX 30-NOV-1999; 99WO-US28313.  
XX 01-DEC-1999; 99WO-US28301.  
XX 01-DEC-1999; 99WO-US28634.  
XX 05-JAN-2000; 2000WO-US00219.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KT,  
PI Marsters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;  
XX Matanabe CK, Wood WI;

DR WPI; 2002-205567/26.  
DR N-PSDB; ABR40255.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for  
PT treating benign or malignant tumours, leukaemias and lymphoid  
PT malignancies, inflammatory, angiogenic and immunologic disorders -  
XX  
PS Claim 61: Fig 4; 302pp; English.

XX The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The  
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
CC macropthalgal, stromal and blastocoeil disorders, inflammatory, immune  
CC and angiogenic disorders. The polynucleotide sequences are also  
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO  
CC polypeptides of the invention.

SQ Sequence 249 AA:

Query Match 87.8%; Score 1268; DB 23; Length 249;  
Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRSQRRRRGRGEGTALVPLALGCLALACGLLAVVSLGSRASLSAQEPAGEEL 95  
DB 1 MAARRSQRRRRGRGEGTALVPLALGCLALACGLLAVVSLGSRASLSAQEPAGEEL 60  
QY 96 VAEEDDPSELNPTQESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 155  
DB 61 VAEEDDPSELNPTQESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 120  
QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGFEITYRAGLYLYLCQVHDEKAVYLLKD 215  
DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGFEITYRAGLYLYLCQVHDEKAVYLLKD 180  
QY 216 LTVDGVALRCLEESFATRAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHKLKAAAPFL 275  
DB 181 LTVDGVALRCLEESFATRAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHKLKAAAPFL 240  
QY 276 TYFGLEFQVH 284  
DB 241 TYFGLEFQVH 249

# RESULT 6

AAW29745 standard; Protein: 249 AA.

XX AAW29745;  
XX 27-OCT-1998 (first entry)  
DE TNF related endothelium proliferative agent protein.  
KW TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
OS Homo sapiens.

XX MO9835061-A2.  
XX 13-AUG-1998.  
XX 12-FEB-1998; 98WO-US02859.  
XX 10-FEB-1998; 98US-0021706.  
XX 12-FEB-1997; 97US-0798692.

XX (ABBO ) ABBOTT LAB.  
XX WALLEY SR;

DR WPI: 1998-447255/38.  
 DR N-PSDB: AAV47613.  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 treatment of autoimmune disease; tumours and inflammation  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 SQ Sequence 249 AA:  
 Query Match 87.6%; Score 1265; DB 19; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 36 MAARRSORRRGRGEGTALVPLALGLALACLGILLAVSLGSRASLSAOPAOEEL 95  
 DB 1 MAARRSQKRRGRGEGTALVPLALGLALACLGILLAVSLGSRASLSAOPAOEEL 60  
 OY 96 VAEEODDPSSELNQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGD 155  
 DB 61 VAEEODDPSSELNQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGD 120  
 OY 156 GAAGVDTGTVSGWEARINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYTKLD 215  
 DB 121 GAAGVDTGTVSGWEARINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYTKLD 180  
 OY 216 LTVDGVALRCLLEEFSSATAASSIGPOLRLCOVSGLLALRPGSSLRITLTPWALKKAAPFL 275  
 DB 181 LTVDGVALRCLLEEFSSATAASSIGPOLRLCOVSGLLALRPGSSLRITLTPWALKKAAPFL 240  
 OY 276 TYRGLFOVH 284  
 DB 241 TYRGLFOVH 249  
 RESULT 7  
 ID AAE00891 standard; Protein: 249 AA.  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 KM Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KM TREPA: TNF related endothelium proliferative agent; tumour; metastasis;  
 KM grafting; vulneryary.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 XX /label= Extracellular\_domain  
 XX US6207642-B1.  
 XX

PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB: AAD04350.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 SQ Sequence 249 AA:  
 Query Match 87.6%; Score 1265; DB 22; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 36 MAARRSORRRGRGEGTALVPLALGLALACLGILLAVSLGSRASLSAOPAOEEL 95  
 DB 1 MAARRSQKRRGRGEGTALVPLALGLALACLGILLAVSLGSRASLSAOPAOEEL 60  
 OY 96 VAEEODDPSSELNQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGD 155  
 DB 61 VAEEODDPSSELNQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGD 120  
 OY 156 GAAGVDTGTVSGWEARINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYTKLD 215  
 DB 121 GAAGVDTGTVSGWEARINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYTKLD 180  
 OY 216 LTVDGVALRCLLEEFSSATAASSIGPOLRLCOVSGLLALRPGSSLRITLTPWALKKAAPFL 275  
 DB 181 LTVDGVALRCLLEEFSSATAASSIGPOLRLCOVSGLLALRPGSSLRITLTPWALKKAAPFL 240  
 OY 276 TYRGLFOVH 284  
 DB 241 TYRGLFOVH 249  
 RESULT 8  
 ID AAU03499 standard; Protein: 273 AA.  
 AC AAU03499;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 KM TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis;  
 KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KM

preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 KM fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 19-DEC-2000; 2000WO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMM) IMMUNEX CORP.  
 XX  
 PI Wiley SR;  
 PI  
 DR WPI: 2001-417975/44.  
 DR N-PSDB; AAS03964.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor.  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 XX  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-LZ-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAK agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant tumours and preneoplastic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 CC  
 XX  
 SQ Sequence 273 AA;  
 Query Match 73.8%; Score 1066; DB 22; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

78 SLGSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPPLNRLVPRRSPKGRKTR 137  
 DB SLGSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPPLNRLVPRRSPKGRKTR 126  
 QY 138 ARRAIAAHYEVHPRGQDGAQAGVDGTVSGWEARINSSPLRNROIGEFIVTRAGLYY 197  
 DB 127 ARRAIAAHYEVHPRGQDGAQAGVDGTVSGWEARINSSPLRNROIGEFIVTRAGLYY 186  
 QY 198 LYCOVHDEGKAVYIKLIDLVGVIALRCLEFSTAASSIGPQRLCQVSGLLARPGS 257  
 DB 187 LYCOVHDEGKAVYIKLIDLVGVIALRCLEFSTAASSIGPQRLCQVSGLLARPGS 246  
 QY 258 SLRIRTLPMWHLKAPFLTYFGLFOVH 284

DB 247 SLRIRTLPMWHLKAPFLTYFGLFOVH 273  
 RESULT 9  
 AAW93590  
 ID AAW93590 standard; Protein; 208 AA.  
 XX  
 AC AAW93590;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNFR3 protein.  
 XX  
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KM developmental abnormality; gestational abnormality; prostate cancer;  
 KM APO6; APO8; APO9; TNFR-1; TNFR-3; diagnosis; treatment; therapy; disease;  
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KM apoptosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIT) UNIT WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 PI  
 DR WPI: 1999-205191/17.  
 DR N-PSDB; AAX23424.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PT  
 PS Claim 40; Fig 13A; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNFR1 and TNFR3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO4 polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX  
 SQ Sequence 208 AA;  
 Query Match 73.5%; Score 1062; DB 20; Length 208;  
 Best Local Similarity 99.5%; Pred. No. 2.6e-93;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

77 VSLSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPPLNRLVPRRSPKGRKTR 136  
 VSLSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPPLNRLVPRRSPKGRKTR 136



Db 1 VSLGSRASLSAQPAPQAEELVAEEDDPSELNPQTEESQDPAPFLNRLNVRPARRSAPGRKT 60  
 QY 137 RARRAAIAHYEVRPRPGDGAAGVDTGYSGEARINSSPLRTNRQIGETIVTRAGLY 196  
 Db 61 RARRAAIAHYEVRPRPGDGAAGVDTGYSGEARINSSPLRTNRQIGETIVTRAGLY 120  
 QY 197 YLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATTAASSIGPOLRCQVSGLLALRPG 256  
 Db 121 YLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATTAASSIGPOLRCQVSGLLALRPG 180  
 QY 257 SLSLRITLPMWHLKAPFLTYFGLFQVH 284  
 Db 181 SLSLRITLPMWHLKAPFLTYFGLFQVH 208

RESULT 10  
 AAM47524  
 ID AAM47524 standard; Protein; 225 AA.  
 AC AAM47524;  
 XX  
 DT 21-JUL-1998 (first entry)  
 XX  
 DE Mus musculus tumour necrosis factor related ligand (TRELL).  
 XX  
 KW TRELL: tumour necrosis factor related ligand; tnf; treatment:  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..21  
 FT /note= "hydrophobic, transmembrane domain"  
 XX  
 PN WO9805783-A1.  
 XX  
 PD 12-FEB-1998.  
 XX  
 PE 07-AUG-1997; 97WO-US13945.  
 XX  
 PR 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA (URGE-) UNIV GENEVA FACULTY MEDICINE.  
 XX  
 PI Browning JL, Chicheportliche Y;  
 XX  
 DR WPI: 1998-145619/13.  
 DR N-PSDB; AAV18599.  
 XX  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 48-50; 69pp; English.

The sequence is that of mouse tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon- gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. It's coding sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective

CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of probe probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 XX  
 SQ Sequence 225 AA;  
 Query Match 70.6%; Score 1020; DB 19; Length 225;  
 Best Local Similarity 88.8%; Pred. No. 2,9e-89;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 61 LGLGLAACIGLLAVVSLGSRASLSAQPAPQAEELVAEEDDPSELNPQTEESQDPAPFL 120  
 Db 2 LGLGLAACIGLLAVVSLGSRASLSAQPAPQAEELVAEEDDPSELNPQTEESQDPAPFL 61  
 QY 121 NRVLRPRRSAPGRKTRARRAIAHYEVRPRPGDGAAGVDTGYSGEARINSSPLR 180  
 Db 62 EQLVPRRSAPGRKTRARRAIAHYEVRPRPGDGAAGVDTGYSGEARINSSPLR 121  
 QY 181 YNRQIGETIVTRAGLYYLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATTAASSIGP 240  
 Db 122 YNRQIGETIVTRAGLYYLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATTAASSIGP 181  
 QY 241 QLRLQVSGLLALRPGSSLSLRITLPMWHLKAPFLTYFGLFQVH 284  
 Db 182 QLRLQVSGLLALRPGSSLSLRITLPMWHLKAPFLTYFGLFQVH 225

RESULT 11  
 AAB07527  
 ID AAB07527 standard; protein; 225 AA.  
 XX  
 AC AAB07527;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200042073-A1.  
 XX  
 PD 20-JUL-2000.  
 XX  
 PE 14-JAN-2000; 2000WO-US01044.  
 XX  
 PR 15-JAN-1999; 99US-0116168.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA Rennert P;  
 PI  
 XX  
 DR WPI: 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.

The present sequence represents a TWEAK protein. The specification describes a method for preventing or treating an immunological disorder and/or inhibiting an immune response in an animal. The method comprises administering a TWEAK blocking agent. The method may be used for preventing and treating immune disorders associated with inappropriate expression and/or activity of TWEAK. These disorders include autoimmune diseases, acute and chronic inflammation, organ transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell malignancies, septic and other forms of shock, loss of immune

CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 CC  
 XX Sequence 225 AA;  
 SO Query Match 70.6%; Score 1020; DB 21; Length 225;  
 Best Local Similarity 88.8%; Pred. No. 2.9e-89;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 61 LGIGLALACGLLAVVSLASRAESLAEPAEELVAEEDQSEELNPTQESODPAPFL 120  
 DB 2 LSLGLALACGLLAVVSLASRAESLAEPAEELVAEEDQSEELNPTQESODPAPFL 61  
 QY 121 NRLVPRRSAPKGRKTRARRAIAHYEVHPRPGDGAQAGVGTSGWEARINSSPLR 180  
 DB 62 EQLVPRRSAPKGRKTRARRAIAHYEVHPRPGDGAQAGVGTSGWEARINSSPLR 121  
 QY 181 YNFOIGFITYRAGLYLYCOVHFDEGKAVYLYKIDLVGVLAIRCLSEFSATPAASSIGP 240  
 DB 122 YDROIGFITYRAGLYLYCOVHFDEGKAVYLYKIDLVGVLAIRCLSEFSATPAASSIGP 181  
 QY 241 QLRICQVSGLLAPRPGSSLRIRTLPMWHLKAPFLTYGFLFOVH 284  
 DB 182 QLRICQVSGLLAPRPGSSLRIRTLPMWHLKAPFLTYGFLFOVH 225  
 RESULT 12  
 AAM93591  
 ID AAM93591 standard; Protein; 211 AA.  
 AC AAM93591;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Mouse TNRL3 protein.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse.  
 XX  
 OS Mus sp.  
 XX  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR N-PSDB: AAX23425.  
 XX  
 PT New Tumour Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumour Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC molecule. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of

CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developing antitumor or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 SO Sequence 211 AA;  
 QY Query Match 67.0%; Score 968; DB 20; Length 211;  
 Best Local Similarity 89.1%; Pred. No. 2.5e-84;  
 Matches 188; Conservative 9; Mismatches 14; Indels 0; Gaps 0;  
 QY 74 LAVVSLGSRASLSAEOEELVAEEDQSEELNPTQESODPAPFLNRLVPRRSAPK 133  
 DB 1 LVVVSLSGSMATLSAQESQSEELVAEDRREPELNPOTESQDVYVFLQVLRPRRSAPK 60  
 QY 134 RKTARRAIAHYEVHPRPGDGAQAGVGTSGWEARINSSPLRYNROIGFITYTRA 193  
 DB 61 RKTARRAIAHYEVHPRPGDGAQAGVGTSGWEARINSSPLRYNROIGFITYTRA 120  
 QY 194 GLYLYCOVHFDEGKAVYLYKIDLVGVLAIRCLSEFSATPAASSIGPRLCOVSGLLAL 253  
 DB 121 GLYLYCOVHFDEGKAVYLYKIDLVGVLAIRCLSEFSATPAASSIGPRLCOVSGLLAL 180  
 QY 254 RRGSSLRIRTLPMWHLKAPFLTYGFLFOVH 284  
 DB 181 RRGSSLRIRTLPMWHLKAPFLTYGFLFOVH 211  
 RESULT 13  
 AAM29746  
 ID AAM29746 standard; Protein; 189 AA.  
 AC AAM29746;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE TNF related endothelium proliferative agent protein 2.  
 XX  
 KW TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Willey SR;  
 XX  
 DR WPI: 1998-447255/38.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 125-6; 142pp; English.  
 CC The TNF-related endothelium proliferative agent (TREPA), or its

CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).

XX Sequence 189 AA;

Query Match 54.8%; Score 792; DB 19; Length 189;  
 Best local Similarity 99.3%; Pred. No. 1.4e-67;  
 Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 132 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 191  
 DB 37 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 96  
 QY 192 RAGLYTYLCOVHDEGKAVYLRKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 251  
 DB 97 RAGLYTYLCOVHDEGKAVYLRKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 156  
 QY 252 ALRPGSSLRIRITLPMHAKAPFLTYFGLEFQVH 284  
 DB 157 ALRPGSSLRIRITLPMHAKAPFLTYFGLEFQVH 189

RESULT 14

AAE00892 AAE00892 standard; Protein; 189 AA.

XX AAE00892;

DT 04-JUL-2001 (first entry)

DE Human UL4flag TREPA soluble construct.

KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;

KM TREPA; TNF related endothelium proliferative agent; metastasis; grafting;

KW vulnery; HUVEC; human umbilical vein endothelial cell; UL4flag.

OS Homo sapiens.

PN US6207642-B1.

PD 27-MAR-2001.

PF 26-JUN-1998; 98US-0105343.

PR 12-FEB-1997; 97US-0798692.

PR 10-FEB-1998; 98US-0021706.

PA (ABBO ) ABBOTT LAB.

PI Wiley SR;

DR WPI; 2001-280760/29.

XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX Example 2; Column 75-78; 53pp; English.  
 CC The present invention relates to extracellular signal molecules,

CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts. The present  
 CC amino acid sequence is human UL4flag TREPA soluble construct. This  
 CC sequence which is a biologically active molecule is capable of inducing  
 CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.

XX Sequence 189 AA;

Query Match 54.8%; Score 792; DB 22; Length 189;  
 Best local Similarity 99.3%; Pred. No. 1.4e-67;  
 Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 132 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 191  
 DB 37 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 96  
 QY 192 RAGLYTYLCOVHDEGKAVYLRKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 251  
 DB 97 RAGLYTYLCOVHDEGKAVYLRKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 156  
 QY 252 ALRPGSSLRIRITLPMHAKAPFLTYFGLEFQVH 284  
 DB 157 ALRPGSSLRIRITLPMHAKAPFLTYFGLEFQVH 189

RESULT 15

AAE00895 AAE00895 standard; Protein; 146 AA.

XX AAE00895;

DT 04-JUL-2001 (first entry)

DE Human TREPA (TNF related endothelium proliferative agent) fragment.

KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;

KM TREPA; TNF related endothelium proliferative agent; tumour; metastasis;

KW grafting; vulnery.

OS Homo sapiens.

PN US6207642-B1.

PD 27-MAR-2001.

PF 26-JUN-1998; 98US-0105343.

PR 12-FEB-1997; 97US-0798692.

PR 10-FEB-1998; 98US-0021706.

PA (ABBO ) ABBOTT LAB.

PI Wiley SR;

DR WPI; 2001-280760/29.

XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX Example 14; Fig 1; 53pp; English.  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted

CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is human TREPA fragment.  
 XX

Sequence 146 AA;

Query Match 52.7%; Score 761; DB 22; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-65;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHVEVHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEIFYTRAGLYYL 198  
 |||  
 DB 1 RRAIAHVEVHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEIFYTRAGLYYL 60  
 OY 199 YCOVHDEEGKAVYLYKIDLVDCYLALRCLEBSATSIGPOLRLCOVSGLALRPSS 258  
 |||  
 DB 61 YCOVHDEEGKAVYLYKIDLVDCYLALRCLEBSATSIGPOLRLCOVSGLALRPSS 120  
 OY 259 LRIRTLPMWHLKAPFLTYFGLFOVH 284  
 |||  
 DB 121 LRIRTLPMWHLKAPFLTYFGLFOVH 146

Search completed: May 8, 2003, 02:19:42  
 Job time : 37.2672 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 05:32:15 ; Search time 58 Seconds  
(without alignments)  
1008.920 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPLTFYGLFQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.5	111	16	08x4j8
2	9	3.2	142	16	0981j0
3	8	2.8	51	4	09BX52
4	8	2.8	143	17	09HST7
5	8	2.8	151	10	09SDI1
6	8	2.8	184	5	09VY70
7	8	2.8	197	16	09WTU0
8	8	2.8	199	4	09BU11
9	8	2.8	211	5	08SXH4
10	8	2.8	278	16	09ZG99
11	8	2.8	279	11	09D378
12	8	2.8	279	11	09CP88
13	8	2.8	306	5	08SZB8
14	8	2.8	339	17	058554
15	8	2.8	342	4	099908
16	8	2.8	343	4	09BPV2

17	8	2.8	370	17	08RTU8	08tut8 methanopyru
18	8	2.8	372	16	053860	053860 mycobacteri
19	8	2.8	377	10	09KXT3	09kxt3 streptomyce
20	8	2.8	387	10	09SHD8	09shd8 arabidopsis
21	8	2.8	397	16	08U820	08u820 agrobacteri
22	8	2.8	435	16	08XS70	08xs70 ralsionia s
23	8	2.8	443	4	09UFM5	09ufm5 homo sapien
24	8	2.8	454	5	09VNP0	09vnp0 drosophila
25	8	2.8	465	16	08ZCV8	08zcv8 yersinia pe
26	8	2.8	471	16	09KRI8	09kri8 delnoccocus
27	8	2.8	472	4	096N66	096n66 homo sapien
28	8	2.8	473	11	09CY76	09cy76 mus musculu
29	8	2.8	473	11	08RIJ9	08rij9 mus musculu
30	8	2.8	522	10	09LTW0	09ltw0 oryza sativ
31	8	2.8	564	12	09YTU9	09ytu9 influenzavi
32	8	2.8	568	12	08OPL0	08opl0 influenza a
33	8	2.8	656	5	09N8H2	09n8h2 trypanosoma
34	8	2.8	1523	5	09N8U8	09n8u8 trypanosoma
35	7	2.5	24	5	P83215	P83215 octopus vul
36	7	2.5	30	5	P83217	P83217 octopus vul
37	7	2.5	32	6	09BEF6	09bef6 capra hircu
38	7	2.5	35	2	09ZG35	09zg35 chlamydia t
39	7	2.5	54	4	016193	016193 homo sapien
40	7	2.5	67	2	08RO84	08ro84 serratia ma
41	7	2.5	69	5	09VMK9	09vmk9 drosophila
42	7	2.5	86	16	09AAG4	09aag4 caulobacter
43	7	2.5	94	4	09P074	09p074 homo sapien
44	7	2.5	99	2	P97199	P97199 escherichia
45	7	2.5	100	5	08WQ88	08wq88 euryymna sc

## ALIGNMENTS

## RESULT 1

08x4j8 PRELIMINARY; PRT; 111 AA.  
ID AC 08X4J8:  
AC 08X4J8:  
DT 01-MAR-2002 (TRMBLrel. 20, Created)  
DT 01-MAR-2002 (TRMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TRMBLrel. 21, Last annotation update)  
DE Hypothetical protein 23516.  
GN 23516.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobbeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.: "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.", Nature 409:529-533(2001).  
RL EMBL: AE005458; AAC57389.1; -.  
DR InterPro: IPR000620; DUF6.  
DR Pfam: PF00892; DUF6.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 111 AA; 12165 MW; 7CEFC93D786CD759 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;

Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

63 LGALACIGL 72  
40 LGALACIGL 49

RESULT 2  
 ID 0981J0 PRELIMINARY: PRT; 142 AA.  
 AC 0981J0:  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE Probable transport protein.  
 GN MUR2380.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Watanabe T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabeta S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti."  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP002999; BAB49526.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 142 AA; 14864 MW; 0DCA7842CB5A56F CRC64;

Query Match 3.2%; Score 9; DB 16; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 TALLVPLAL 61  
 |||||  
 DB 122 TALLVPLAL 130

RESULT 3  
 ID 09BX52 PRELIMINARY: PRT; 51 AA.  
 AC 09BX52:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE DJ831C21.1 (Novel protein similar to gamma-glutamyl  
 DE transpeptidase-related protein (GGF-Rel)) (Fragment).  
 GN DJ831C21.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Cory N.;  
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133466; CAC34607.1; -.  
 FT NON\_TER 51 51  
 SQ SEQUENCE 51 AA; 5124 MW; 31980CCA0451EF8 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 DB 7 LGIGLALA 14

RESULT 4  
 ID 09HST7 PRELIMINARY: PRT; 143 AA.

AC 09HST7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Vng0080H.  
 GN VNG0080H.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20504483; PubMed=11016950;  
 RX Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Bailga N.S., Thorsson V., Shroogna J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,  
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AE004976; AAG18715.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 143 AA; 15648 MW; 45466E6328EF3468 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 DB 55 LGIGLALA 62

RESULT 5  
 ID 09SD11 PRELIMINARY: PRT; 151 AA.  
 AC 09SD11:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical protein (OSJNB0036E02.6 protein) (B1085F09.2  
 DE protein).  
 GN B1085F09.2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P003H10.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:OSJNB0036E02.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:B1085F09.";  
 DL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP000815; BAA87834.1; -.

DR	EML: AP002862; BAB17732.1; -	Score 8; DB 10; Length 151;
DR	EML: AP003103; BAB44106.1; -	Best Local Similarity 100.0%; Pred. No. 15;
SO	SEQUENCE 151 AA; 16632 MW; EC68451ECA2BD71D CRC64;	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy	43 RRRGRGCE 50	
Db	131 RRRGRGCE 138	
RESULT 6		
Q9VW70	PRELIMINARY; PRT; 184 AA.	
AC	Q9VW70;	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
OC	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
GN	CG13033 protein.	
DE	CG13033.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RP	1]	
RC	SEQUENCE FROM N.A.	
RX	STRAIN-BERKELEY;	
RA	MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA	Georgiadis S.E., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Anton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,	
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Adair J.F., Agayani A., An H.-J., Andrews-Plankkoch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beehon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,	
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	
RA	Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA	Gjoeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,	
RA	Jaisl M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Kethum K.A.,	
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	
RA	Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,	
RA	Mont S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,	
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Rennett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA	Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,	
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,	
RA	Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	
RT	"The genome sequence of Drosophila melanogaster."	
RL	Science 287:2185-2195(2000).	
DR	EML: AEO03527; AAF9452.1; -	
DR	PlayBase; FBgn0036638; CG13033	
DR	InterPro; IPR004019; YLP_motif.	
DR	Pfam; Pf02757; YLP_5.	

```

SO SEQUENCE 184 AA; 21022 MW; 31976AE30DDC447 CRC64;
Query Match 2.8%; Score 8; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLNAVSL 79
Db 16 LLNAVSL 23

RESULT 7
O9WYUO
ID O9WYUO PRELIMINARY; PRT; 197 AA.
AC O9WYUO;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein TM0469.
GN TM0469.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM DSM 3109;
RC SPRAIN-MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hatt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Ueberback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001724; AAD35553.1; -.
DR TIGR; TM0469; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 197 AA; 22919 MW; 41E2C8E3C09180EC CRC64;

OY 2 SLDPFIS 9
Db 135 SLDPFIS 142

RESULT 8
O9BUTI
ID O9BUTI PRELIMINARY; PRT; 199 AA.
AC O9BUTI;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 22.8 kDa protein.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strauberg R.;
RL Submitted (Feb.-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002598; AAH02598.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 22750 MW; E5DCA747E8BF06B CRC64;

```

Query Match 2.8%; Score 8; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLAL 67  
DB 167 ALGLAL 174

## RESULT 9

O8SXH4 PRELIMINARY; PRT; 211 AA.

AC O8SXH4; 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DE RE0345p.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;

SEQUENCE FROM N.A.

RA STRAIN-BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Pargas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celisner S.,  
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY089634; AAL90372.1;  
SQ SEQUENCE 211 AA; 23780 MW; 82FF4983E91F510A CRC64;

Query Match 2.8%; Score 8; DB 5; Length 211;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVSL 79  
DB 9 LLLAVSL 16

## RESULT 10

O9ZG99 PRELIMINARY; PRT; 278 AA.

AC O9ZG99; 01-MAY-1999 (TREMblrel. 10, Created)  
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Transmembrane protein AMPE.  
GN AMPE OR PA4521.

OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.

NCBI\_TaxID=287;

SEQUENCE FROM N.A.

RA STRAIN-ATCC 15692 / PA01;  
RA Langae Y.T., Davis M., Huletsky A.;  
RT "An amp gene in Pseudomonas aeruginosa encodes a negative regulator  
of amp beta-lactamase expression."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

RA STRAIN-ATCC 15692 / PA01;  
RA MEDLINE-20437337; PubMed-10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrone M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen";

RL Nature 406:959-964(2000).  
DR EMBL: AF082575; AAC98784.1; -;  
DR EMBL: AF004866; AAG07909.1; -;  
KW Transmembrane; Complete proteome.  
SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CFEF CRC64;

Query Match 2.8%; Score 8; DB 16; Length 278;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 77  
DB 47 LGLLAV 54

## RESULT 11

O9D378 PRELIMINARY; PRT; 279 AA.

AC O9D378; 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DE 5730494G16R1k protein.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
SEQUENCE FROM N.A.

RA STRAIN-C57BL/6J; TISSUE-MEDULLA OBLONGATA;  
RX MEDLINE-21085660; PubMed-11217851;  
RA Kawai J., Shingawa A., Shidata K., Yoshino M., Itch M., Ishii Y.,  
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto J.,  
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Felschmann W., Gaasterland T., Gissi C., King B., Kocher H.,  
RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombaris P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shidata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
DR EMBL: AK018250; BAB31133.1; -;  
DR MGD: MGI:1913897; 5730494G16R1k.  
DR InterPro: IPR002190; MAGE.  
DR Pfam: PF01454; MAGE; 1.  
SQ SEQUENCE 279 AA; 31474 MW; 5E243590A99F15F0 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 80 GSRASLSA 87  
DB 41 GSRASLSA 48

## RESULT 12

O9CPR8 PRELIMINARY; PRT; 279 AA.



```

AC 09CPR8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 5730494G16RIK protein (MAGE-g1).
GN 5730494G16RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO, AND EMBRYONIC STEM CELLS.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Komdeurs P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Williams L.,
RA Wyshniew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Augier P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;
RT "Ten new murine members of the MAGE gene family.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBD databases.
DR EMBL; AK017727; BAB30899.1; -
DR EMBL; AK010294; BAB26830.1; -
DR MGI; MGI:1913897; 5730494G16RIK.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE.1.
SQ SEQUENCE 279 AA; 31460 MW; FE2435919BD63160 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 80 GSRASLSA 87
DB 41 GSRASLSA 48

RESULT 13
OBSZB8 PRELIMINARY; PRT; 306 AA.
AC 08SZB8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE RE07882P.
GN CG1169.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;

```

```

RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dyesnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Munoz J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBD databases.
RX EMBL; AY070982; AAL4604.1; -
SQ SEQUENCE 306 AA; 34083 MW; 32B69371475A48F9 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 306;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 LVPALGL 63
DB 170 LVPALGL 177

RESULT 14
OBSZB8 PRELIMINARY; PRT; 339 AA.
AC 058554;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein PH0824.
GN PH0824.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Hatake Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai K., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohiku Y.,
RA Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000003; BAA2917.1; -
DR InterPro; IPR002657; BileAc_Na_smptr.
DR Pfam; PF01758; SBF.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 339 AA; 37228 MW; E91697D5C8C3705F CRC64;

Query Match 2.8%; Score 8; DB 17; Length 339;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGILLAV 77
DB 111 LGILLAV 118

RESULT 15
OBSZB8 PRELIMINARY; PRT; 342 AA.
AC 099908;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE BBL protein.
GN BBL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96273128; PubMed-8702217;
RA Fukunaga-Johnson N., Lee S.W., Liebert M., Grossman H.B.;
RT "Molecular analysis of a gene, BBI, overexpressed in bladder and
RL breast carcinoma."
RL Anticancer Res. 16:1085-1090(1996).
DR EMBL: S82470; AAB37433.1; -.
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
SQ SEQUENCE 342 AA; 38163 MW; 2B479EABCFB191C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLAL 67
Db 310 ALGLGLAL 317

RESULT 16
Q9BPV2 PRELIMINARY; PRT; 343 AA.
ID Q9BPV2
AC Q9BPV2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:4221) (protein for MGC:2099).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, AND COLON;
RA Strausberg R.;
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC003164; AA003164.1; -.
DR EMBL: BC002512; AA002512.1; -.
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
SQ SEQUENCE 343 AA; 38727 MW; F71E7DBF74BD9BB7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLAL 67
Db 311 ALGLGLAL 318

RESULT 17
Q8TUU8 PRELIMINARY; PRT; 370 AA.
ID Q8TUU8
AC Q8TUU8;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Permease subunit of a ABC-type transport system involved in
DE lipoprotein release.
GN MK1655.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE-21927647; PubMed-11930014;
RA Slesarev A.I., Mezhevaeva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,

```

```

RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malikh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL: AE010455; AA02868.1; -.
RW Complete proteome.
SQ SEQUENCE 370 AA; 39411 MW; B07662EA1E5A644E CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGGLALAL 68
Db 336 LGGLALAL 343

RESULT 18
O53860 PRELIMINARY; PRT; 372 AA.
ID O53860
AC O53860;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein cysM3.
GN CYSM3 OR RV0848 OR MT0043. 41.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,
RA Hovnsdy T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sutcliffe J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL: AL022004; CAA17654.1; -.
DR HSP: P35520; 1JBO.
DR TubercuList; RV0848; -.
DR InterPro: IPR001926; B6_enzyme_beta.
DR Pfam: PF00291; PALP; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 372 AA; 40118 MW; 927386B81DF5F8C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGGLALAL 68
Db 96 LGGLALAL 103

RESULT 19
O9KXT3 PRELIMINARY; PRT; 377 AA.
ID O9KXT3
AC O9KXT3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC05682 OR SC5H4.06C.
OS Streptomyces coelicolor.

```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_Taxid=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M45;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Godle A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Metzger A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL355913; CAB91118.1;  
 SQ SEQUENCE 377 AA; 37614 MW; A35DA0437F044C64 CRC64;  
 Query Match 2.8%; Score 8; DB 16; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 59 LALGGLA 66  
 |||||  
 DB 327 LALGGLA 334  
 RESULT 20  
 O9SHD8 PRELIMINARY; PRT: 387 AA.  
 ID O9SHD8;  
 AC O9SHD8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created).  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE At2g45000 protein (At2g45000/t14p1.20) (Hypothetical 40.6 kDa  
 DE protein).  
 CN AT2G45000 OR AT2G45005.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentto M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RA Bowser L., Carinci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,

RA Ecker J.R.;  
 RT "Arabidopsis cDNA clones.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
 RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shinozaki K., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis Full Length cDNA Clones.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007659; AAD32835.1;  
 DR EMBL: AY074646; AAL69462.1;  
 DR EMBL: AY080620; AAL6303.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 387 AA; 40584 MW; AF6C6B3BAC9BP69A CRC64;  
 Query Match 2.8%; Score 8; DB 10; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 93 EEIVAEED 100  
 |||||  
 DB 370 EEIVAEED 377  
 RESULT 21  
 O8U820 PRELIMINARY; PRT: 397 AA.  
 ID O8U820;  
 AC O8U820;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Atu3948.  
 GN AtU3948 OR AGR L.1808.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_Taxid=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Moks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.";  
 RL Science 294:2317-2323(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Quirello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Hounzel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Martelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Seear C., Strub G.,  
 RA Gioe C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE009325; AAL44750.1; ALT\_INIT.  
 DR EMBL: AE008289; AAK89478.1;  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 397 AA; 41708 MW; 700748E32A46A86 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 397;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258

DB 341 LALRPGSS 348

RESULT 22

08X570

AC 08X570

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DE Probable transmembrane protein.

GN RSP0611 OR RS03756.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Plasmid megaplasmid.

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI\_TaxID=305;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN-GM11000;

RX MEDLINE-21681879; PubMed-11823852;

RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

RA Arlet M., Billault A., Brotlier P., Camus J.C., Catolico L.,

RA Chaudier M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,

RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,

RA Sigler P., Thebault P., Whalen M., Wincker P., Levy M.,

RA Weissenbach J., Boucher C.A.,

RT "Genome sequence of the plant pathogen Ralstonia solanacearum";

RL Nature 415:497-502(2002).

DR EMBL: AL646079; CAD17762.1; -

DR InterPro: PS00469; NDP\_KINASES; UNKNOWN\_1.

DR PROSITE: PS00469; NDP\_KINASES; UNKNOWN\_1.

DR Plasmid: Complete proteome.

SO SEQUENCE 435 AA; 47048 MW; CCD859D9C54DD85A CRC64;

Query Match 2.8%; Score 8; DB 16; Length 435;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDGYLAL 224

DB 358 LVDGYLAL 365

RESULT 23

09UPM5

AC 09UPM5

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE Hypothetical 47.4 kDa protein..

GN DREFP5660011.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RP [1]

RP SEQUENCE FROM N.A.

RC TSSDE-KIDNEY;

RA Blum H., Bauersachs S., Mewes H.W., Gaassenhuber J., Wiemann S.,

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL117144; CAB55910.1; -

DR MEROPS: T03.002; -

DR InterPro: IPR00101; Gglutnspptase.

DR Pfam: PF01019; G\_glu\_transpt; 1.

DR PRINTS: PR01210; GGTTRANSPASE.

KW Hypothetical protein.

SQ SEQUENCE 443 AA; 47446 MW; 3D75E9DB08265971 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 443;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68

DB 14 LGGLALA 21

RESULT 24

09VNP0

AC 09VNP0

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE CG1169 protein.

GN CG1169.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle J., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferriz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassenaar D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).

DR EMBL: AE003600; AAF51889.1; -

DR FLYBase: FBgn0037428; CG1169.

SO SEQUENCE 454 AA; 51320 MW; A75AABD97E716573 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 454;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 LVPALGL 63  
Db 170 LVPALGL 177

## RESULT 25

08ZCV8 PRELIMINARY; PRT; 465 AA.  
AC 08ZCV8;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative membrane protein.  
GN YECB OR YPO2850.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
NCBI\_TaxID=632;  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CO-92 / BIOVAR ORIENTALIS;  
MEDLINE=21470413; PubMed=11586360;

RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Holden M.T.G.,  
RA Prentice M.B., Sebalthia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,  
RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
RL Nature 413:523-527(2001).  
DR EMBL: AJ14154; CAC92102.1; -  
DR InterPro: IPR003662; sub transporter.  
DR Pfam: PF00083; sugat\_trf\_1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 465 AA; 50176 MW; OCC273F10B83F5ED CRC64;

Query Match 2.8%; Score 8; DB 16; Length 465;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLAVALSL 79  
Db 335 LLAVALSL 342

## RESULT 26

09RR18 PRELIMINARY; PRT; 471 AA.  
AC 09RR18;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Transport protein, putative.  
GN DR2502.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
NCBI\_TaxID=1299;  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Mofatt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1."  
RL Science 286:1571-1577(1999).  
DR EMBL: AE002079; AAF12043.1; -  
DR TIGR; DR2502; -  
KW Complete proteome.  
SQ SEQUENCE 471 AA; 47974 MW; 96B2BEBFE445D27 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 471;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLAL 67  
Db 366 ALGLGLAL 373

## RESULT 27

096N66 PRELIMINARY; PRT; 472 AA.  
AC 096N66;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE CDNA FLJ131346 f1s, clone MESAN1000180, highly similar to BBI-malignant  
DE cell expression-enhanced gene/tumor progression-enhanced gene.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NCBI\_TaxID=9606;  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Taahiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
RA Fukuzumi Y., Fujimori Y., Komiyama T., Irie R.,  
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,  
RA Isozaki T.;  
RT "NEDO human CDNA sequencing project."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK055908; BAB71043.1; -  
DR InterPro: IPR004299; MEOAT\_fam.  
DR Pfam: PF03062; MEOAT; 1.  
SQ SEQUENCE 472 AA; 52774 MW; EA721998043F9EBD CRC64;

Query Match 2.8%; Score 8; DB 4; Length 472;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLAL 67  
Db 440 ALGLGLAL 447

## RESULT 28

09CY76 PRELIMINARY; PRT; 473 AA.  
AC 09CY76;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE 5730589L02RIK protein.  
GN 5730589L02RIK.  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fieschmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,  
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RT Nature 409:685-690(2001).  
 DR EMBL: AK019981; BAB31950.1; -;  
 DR MGD: MGI:1924832; 5730589L02Rik.  
 DR InterPro: IPR004239; MBOAT\_Fam.  
 DR Pfam: PF03062; MBOAT\_1.  
 SO SEQUENCE 473 AA; 53504 MW; CE6F8E93C3D01C4F CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACIGL 72  
 Db 436 LALACIGL 443

RESULT 29  
 O8R1P9 PRELIMINARY; PRT; 473 AA.  
 ID O8R1P9;  
 AC O8R1P9;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DE RIKEN cDNA 5730589L02 gene.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC023417; AAH23417.1; -;  
 SO SEQUENCE 473 AA; 53382 MW; DAA1FEDDA78013EA CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACIGL 72  
 Db 436 LALACIGL 443

RESULT 30  
 O9L1W0 PRELIMINARY; PRT; 522 AA.  
 ID O9L1W0;  
 AC O9L1W0;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE Similar to an Arabidopsis thaliana chromosome BAC genomic

DE sequence.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartodeae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;  
 RT "Oryza sativa PAC P0699E04 genomics sequence, complete sequence.";  
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP01111; BAA90509.1; -;  
 SO SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRGE 50  
 Db 415 RRRGRGE 422

RESULT 31  
 O9YTU9 PRELIMINARY; PRT; 564 AA.  
 ID O9YTU9;  
 AC O9YTU9;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hemagglutinin H5 (fragment).  
 OS Influenzavirus A.  
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza A viruses.  
 OX NCBI\_TaxID=11320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A/DUCK/PODSAM/2216-4/84;  
 RX MEDLINE-99099002; PubMed-9882316;  
 RA Matrosovich M., Zhou N., Kawoka Y., Webster R.;  
 RT "The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties.";  
 RT J. Virol. 73:1146-1155(1999).  
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
 CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.  
 DR EMBL: AF082041; AAD13573.1; -;  
 DR HSSP: P03437; IHTM.  
 DR InterPro: IPR001364; Hemagglutn.  
 DR Pfam: PF00509; Hemagglutinin.1.  
 DR PRINTS: PRO0329; HEMAGGLUTN12.  
 DR ProDom: PD000225; Hemagglutn.1.  
 KW Envelope protein; Glycoprotein; Hemagglutinin.  
 FT NON\_TER 564  
 FT 564  
 SO SEQUENCE 564 AA; 63562 MW; B317179A7E3E6F98 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVSL 79  
 Db 6 LLLAVSL 13

RESULT 32  
 O8QPL0 PRELIMINARY; PRT; 568 AA.  
 ID O8QPL0;  
 AC O8QPL0;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 CN Hemagglutinin (Fragment).  
 OS Influenza A virus (A/Goose/Hong Kong/3014.8/2000(H5N1)).  
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza A viruses; Influenzavirus A.  
 OX NCBI\_Taxid-176675;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A/GOOSE/HONG KONG/3014.8/2000(H5N1);  
 RA MEDLINE-21874832; PubMed-11878904;  
 RA Guan Y., Peiris M., Kong K.F., Dyrting K.C., Ellis T.M., Sit T.,  
 RA Zhang L.J., Shortridge K.F.;  
 RT \*H5N1 Influenza Viruses Isolated from Geese in Southeastern China:  
 RT Evidence for Genetic Reassortment and Interspecies Transmission to  
 RT Ducks.\*;  
 RL Virology 292:16-23(2002).  
 DR EMBL: AY059482; AL31388.1; .  
 FT NON\_TER 568 568  
 SQ SEQUENCE 568 AA; 64281 MW; 0B0ACE034F1769 CRC64;  
  
 Query Match 2.8%; Score 8; DB 12; Length 568;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 72 LLLAVSL 79  
 DB 6 LLLAVSL 13  
  
 RESULT 33  
 ID Q9N8H2 PRELIMINARY; PRT; 656 AA.  
 AC Q9N8H2;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
 DE Possible hypothetical 61.9 kDa protein.  
 CN CHRI.338.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_Taxid-5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FRE927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerard C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL359782; CAB95571.1; .  
 FT NON\_TER 656 656  
 SQ SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;  
  
 Query Match 2.8%; Score 8; DB 5; Length 656;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 93 EELVAED 100  
 DB 454 EELVAED 461  
  
 RESULT 34  
 ID Q9N8U8 PRELIMINARY; PRT; 1523 AA.  
 AC Q9N8U8;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE Possibly hypothetical protein 85.6 kD.  
 CN CHRI.139.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI\_Taxid-5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FRE927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerard C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL359782; CAB95435.1; .  
 FT NON\_TER 1523 1523  
 SQ SEQUENCE 1523 AA; 168322 MW; 07BDC751CDD1E5A CRC64;  
  
 Query Match 2.8%; Score 8; DB 5; Length 1523;  
 Best Local Similarity 100.0%; Pred. No. 1,2e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 218 VDGVALR 225  
 DB 382 VDGVALR 389  
  
 RESULT 35  
 ID P83215 PRELIMINARY; PRT; 24 AA.  
 AC P83215;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)  
 DE Sperm protamine p3 (P03) (Fragment).  
 OS Octopus vulgaris (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurrata; Octopodidae; Octopus.  
 OX NCBI\_Taxid-6645;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE-SPERM;  
 RA Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Kouach M.,  
 RA Ausio J., Kasinsky H.E., Chiva M.;  
 RT \*Chromatin remodelling and protamines during spermiogenesis of Octopus  
 RT vulgaris (Cephalopoda).\*;  
 RL J. Exp. Zool. 0:0-0(2001).  
 CC -1- FUNCTION: PROTEIN SUBSTITUTION FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- MASS SPECTROMETRY: MW=4389; METHOD=ELECTROSPRAY.  
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW DNA condensation; Nuclear protein.  
 FT DOMAIN 1 16 POLY-ARG.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 3381 MW; 308E90ED9D2C99C CRC64;  
  
 Query Match 2.5%; Score 7; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 43 RRRGRG 49  
 DB 11 RRRGRG 17  
  
 RESULT 36  
 ID P83217 PRELIMINARY; PRT; 30 AA.  
 AC P83217;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE Sperm protamine p5 (P05).  
 OS Octopus vulgaris (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurrata; Octopodidae; Octopus.

OX NCB1\_TaxID=6645;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE-SPERM;  
 RA Gimenez-Bonafé P., Ribes E., Buesa C., Sautiere P., Kouach M.,  
 RA Ausio J., Kasinsky H.E., Chiva M.;  
 RT "Chromatin remodeling and protamines during spermatogenesis of Octopus  
 RT vulgaris (Cephalopoda).";  
 RL J. Exp. Zool. 0:0-0(2001).  
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- MASS SPECTROMETRY: MW=3941; METHOD=ELECTROSPRAY.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KM DNA condensation; Nucleic protein.  
 FT DOMAIN 2 POLY-ARG.  
 FT DOMAIN 17 POLY-ARG.  
 FT DOMAIN 26 POLY-ARG.  
 SQ SEQUENCE 30 AA; 3943 MW; 14F1BC7E4D277049 CRC64;

Query Match 2.5%; Score 7; DB 5; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRGRRG 49  
 DB 5 RRGRRG 11

RESULT 37  
 O9BEF6 PRELIMINARY; PRT; 32 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Beta-lactoglobulin precursor (Fragment).  
 GN BETA-LACTOGLOBULIN.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCB1\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA D'Andrea M., Pilla F., Graziano M.;  
 RT "A new polymorphism of goat beta-lactoglobulin proximal promoter  
 RT region.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ292058; CAC27455.1;  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT NON\_TER 32 32  
 SQ SEQUENCE 32 AA; 3372 MW; 0C56BD579B3DC190 CRC64;

Query Match 2.5%; Score 7; DB 6; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALAC 69  
 DB 8 LGLALAC 14

RESULT 38  
 O9ZG35 PRELIMINARY; PRT; 35 AA.  
 ID O9ZG35  
 AC O9ZG35  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE Hypothetical 3.5 kDa protein (Fragment).

OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCB1\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L2 434b;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF087333; AAD4107.1;  
 KW Hypothetical protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 35 35  
 SQ SEQUENCE 35 AA; 3519 MW; 46686B72BAF28F7D CRC64;

Query Match 2.5%; Score 7; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NSSSPLR 180  
 DB 8 NSSSPLR 14

RESULT 39  
 O16193 PRELIMINARY; PRT; 54 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Orf25, to PD-ECGF/TP protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94312438; PubMed=8038210;  
 RX Usuki K., Genez L.J., Wernstedt C., Moren A., Miyazono K.,  
 RA Claesson-Welsh L., Heldin C.H.;  
 RT "Structural properties of 3.0 kb and 3.2 kb transcripts encoding  
 RT platelet-derived endothelial cell growth factor/thymidine  
 RT phosphorylase in A431 cells."  
 RL Biochim. Biophys. Acta 1222:411-414(1994).  
 DR EMBL: S72487; AAD14107.1;  
 SQ SEQUENCE 54 AA; 5808 MW; 827925FBA70CB222 CRC64;

Query Match 2.5%; Score 7; DB 4; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 RPRRSAP 131  
 DB 17 RPRRSAP 23

RESULT 40  
 O8ROS4 PRELIMINARY; PRT; 67 AA.  
 ID O8ROS4  
 AC O8ROS4  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 7.8 kDa protein.  
 GN YHCR.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 OX NCB1\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.



RA Baba M., Midorikawa Y., Nakagawa Y., Fujita M., Matsuyama T.:  
 RT "Serratia marcescens and Escherichia coli genes controlling  
 RT temperature-dependent production of structurally unrelated secondary  
 RT metabolites such as prodigiosin and serrawettin.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB080601; BAB85653.1; -  
 RM Hypothetical protein.  
 SO SEQUENCE 67 AA; 7791 MW; 47B01A87E69AC2A2 CRC64;

Query Match 2.5%; Score 7; DB 2; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LALVPLAL 61  
 Db 21 LALVPLAL 27

## RESULT 41

O9VMM9 PRELIMINARY; PRT; 69 AA.  
 AC O9VMM9; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG11147 protein.  
 GN CG11147.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA George R.A., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Balding D.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Fouts D., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson S.D., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 RA Relibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glids R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).  
 DR EMBL; AE003611; AAF52284.1; -  
 DR Flybase; FBgn0031734; CG11147.  
 SO SEQUENCE 69 AA; 8118 MW; 9728A36DDE94915F CRC64;

Query Match 2.5%; Score 7; DB 5; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRRGR 48  
 Db 54 ORRRGR 60

## RESULT 42

O9A4G4 PRELIMINARY; PRT; 86 AA.  
 AC O9A4G4; 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein CC2870.  
 GN CC2870.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OC NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eissen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,  
 RA Utlarbeck T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).  
 DR EMBL; AE005952; AAK24834.1; -  
 DR TIGR; CC2870; -  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 86 AA; 7615 MW; EEF6383D9A545E CRC64;

Query Match 2.5%; Score 7; DB 16; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGICL 65  
 Db 26 LALGICL 32

## RESULT 43

O9P074 PRELIMINARY; PRT; 94 AA.  
 AC O9P074; 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HSPC308 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BLOOD;  
 RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,  
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;  
 RT "Human partial CDS from cd34+ stem cells.";

RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF161426; AAF28986.1; -  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10874 MW; CDA7EPD97C850842P CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 7; DB 4; Length 94;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 98 EEDDDPS 104  
 |||||  
 Db 11 EEDDDPS 17  
 RESULT 44  
 P97199 PRELIMINARY; PRT; 99 AA.  
 ID P97199;  
 AC P97199;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE Similar to.  
 GN IS5.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251358; Pubmed-9097040;  
 RA Itoh T., Alaba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,  
 Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,  
 Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,  
 Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,  
 Yamamoto Y., Horinouchi T.;  
 RA "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 40,150,000 min region on the linkage map.";  
 RL DNA Res. 3:379-392(1996).  
 DR EMBL: D90837; BAA15812.1; -  
 DR EMBL: D90838; BAA15817.1; -  
 SQ SEQUENCE 99 AA; 10142 MW; 03EBF8DD10E92C0 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 99;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 23 DGGAVRQ 29  
 |||||  
 Db 88 DGGAVRQ 94  
 RESULT 45  
 Q8MOR8 PRELIMINARY; PRT; 100 AA.  
 ID Q8MOR8;  
 AC Q8MOR8;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Lox4 homeodomain protein (Fragment).  
 GN LOX4.  
 OS Euprymna scolopes.  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Sepioida;  
 OC Sepiidae; Euprymna.  
 OX NCBI\_TaxID=6613;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Calleerts P., Lee P.N., Hartmann B., Farfan C.B., Choy D.W.Y.,  
 RA Fischbach K.F., Gehring W.J., de Couet H.G.;  
 RT "HOX genes in the sepioid squid Euprymna scolopes: implications for  
 the evolution of complex body plans.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AT052759; AAL25810.1; -

DR InterPro; IPR001827; Antennapedia.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_repressor.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00025; ANTENNAPEDIA.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR Prodom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; UNKNOWN\_1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 100 AA; 12053 MW; 26C262722E4DC514 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 7; DB 5; Length 100;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 41 SQRRGR 47  
 |||||  
 Db 18 SQRRGR 24  
 Search completed: May 8, 2003, 06:57:36  
 Job time : 62 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 04:39:54 ; Search time 35 Seconds  
(without alignments)

336.551 Million cell updates/sec

Title: US-09-245-198A-4  
Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPFLTYGELFQVH 284

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt.40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	87.7	249	1	TN12_HUMAN
2	32	11.3	225	1	TN12_MOUSE
3	10	3.5	111	1	YFWM_ECOLI
4	9	3.2	733	1	PSAB_ODOSI
5	9	3.2	734	1	PSAB_CYACA
6	8	2.8	179	1	ADHS_GLOXC
7	8	2.8	220	1	Y304_BROME
8	8	2.8	317	1	MSHR_PANTR
9	8	2.8	379	1	FDHB_METJA
10	8	2.8	422	1	ZP3_MESAV
11	8	2.8	576	1	CYDC_HAETN
12	8	2.8	586	1	GCT5_HUMAN
13	7	2.5	49	1	HSP1_SAGIM
14	7	2.5	57	1	HSP1_DIDMA
15	7	2.5	115	1	A62F_DROME
16	7	2.5	118	1	Y151_ECOLI
17	7	2.5	131	1	IL13_MOUSE
18	7	2.5	131	1	IL13_RAT
19	7	2.5	147	1	YK01_PIRHO
20	7	2.5	150	1	TNFC_PIG
21	7	2.5	157	1	RA05_ORYSA
22	7	2.5	161	1	HLPA_ECOLI
23	7	2.5	170	1	Y1M4_CAEEL
24	7	2.5	180	1	LACB_BUBBU
25	7	2.5	180	1	LACB_CAPII
26	7	2.5	180	1	LACB_SHEEP
27	7	2.5	230	1	RNFE_SALTY
28	7	2.5	231	1	RNFE_ECO57
29	7	2.5	231	1	RNFE_ECOLI
30	7	2.5	233	1	RNFC_COXIB
31	7	2.5	233	1	RNFE_YERPE
32	7	2.5	235	1	RNFE_HAETN
33	7	2.5	239	1	TN14_MOUSE

## ALIGNMENTS

ID	TN12_HUMAN	STANDARD	PRT	249 AA.
AC	043508: Q8WU27;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).			
GN	TNFRSF12 OR APO3L OR DR31G.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.			
RC	TISSUE-Tonsil, and fetal liver;			
RX	MEDLINE=98070415; PubMed=9405449;			
RA	Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,			
RA	Hession C., Garcia I., Browning J.L.;			
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."			
RT	J. Biol. Chem. 272:32401-32410(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE-Fetal kidney;			
RC	MEDLINE=98228355; PubMed=9560343;			
RA	Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,			
RA	Ashkenazi A.;			
RT	"Identification of a ligand for the death-domain-containing receptor APO3."			
RT	Curr. Biol. 8:525-528(1998).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE-Tonsil;			
RC	Strausberg R.;			
RA	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.			
RL	[4]			
RN	FUNCTION.			
RP	PubMed=10085077;			
RA	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;			
RT	J. Biol. Chem. 274:8455-8459(1999).			
RL	J. Biol. Chem. 274:8455-8459(1999).			
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.			
CC	-1- SUBUNIT: HOMOTRIMER (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.			
CC	-1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas, skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peritoneal blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.			
CC	-1- PM: The soluble form derives from the membrane form			

CC by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a  
 CC frameshift in position 125.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF030099; AAC51923.1; -;  
 CC EMBL: AF055872; AAC39724.1; -;  
 CC EMBL: BC019047; AAH19047.1; ALT\_FRAME.  
 CC Genew: HGNC:11927; TNFSF12.  
 CC MIM: 602695; -  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF\_1.  
 CC SMART: SM00207; TNF\_1.  
 CC PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC PROSITE: PS00049; TNF\_2; 1.  
 CC KEGG: Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC CHAIN 1 249  
 CC FT CHAIN 1 249  
 CC FT CHAIN 94 249  
 CC FT DOMAIN 1 21  
 CC FT TRANSMEM 22 42  
 CC FT SITE 43 249  
 CC FT CARBOHYD 139 139  
 CC FT SEQUENCE 249 AA; 27216 MW; E660843361C28BBA CRC64;  
 CC  
 CC Query Match 87.7%; Score 249; DB 1; Length 249;  
 CC Best Local Similarity 100.0%; Pred.No.3.2e-231;  
 CC Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 36 MAARRSRRGRGRGEPGTALVPLALGLALACIGLLAVVSGSRASASAEPAQEE 95  
 CC DB 1 MAARRSRRGRGRGEPGTALVPLALGLALACIGLLAVVSGSRASASAEPAQEE 60  
 CC QY 96 VAEEDDPSELNPQTEESQDPAPFLNLRVPRRSAPRGKRTARRAIAHYEYHPRGOD 155  
 CC DB 61 VAEEDDPSELNPQTEESQDPAPFLNLRVPRRSAPRGKRTARRAIAHYEYHPRGOD 120  
 CC QY 156 GAGAGVDCGTSGWEARINSSPLRYNRQIGETVTRAGIYLYCQVHPEGKAAYIKLD 215  
 CC DB 121 GAGAGVDCGTSGWEARINSSPLRYNRQIGETVTRAGIYLYCQVHPEGKAAYIKLD 180  
 CC QY 216 LTVDGVALRCLDEFSATASISLGPRLCQVSGLLARPGSSLRITTLPMHAKAPFL 275  
 CC DB 181 LTVDGVALRCLDEFSATASISLGPRLCQVSGLLARPGSSLRITTLPMHAKAPFL 240  
 CC QY 276 TYFGLEQVH 284  
 CC DB 241 TYFGLEQVH 249  
 CC  
 CC RESULT 2  
 CC TN12\_MOUSE STANDARD; PRT; 225 AA.  
 CC ID TN12\_MOUSE  
 CC AC 054907; Q9CTP2;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak  
 CC inducer of apoptosis) (TWEAK) (Fragment).  
 CC GN TNFSF12.  
 CC OS Mus musculus (Mouse).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 CC [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Peritoneal macrophage;  
 CC RX MEDLINE=98070415; PubMed=9405449;  
 CC RA Chlcheporliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,  
 CC RT Hesson C., Garcia I., Browning J.L.;  
 CC RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that  
 CC weakly induces apoptosis."  
 CC RL J. Biol. Chem. 272:32401-32410(1997).  
 CC [2]  
 CC RP SEQUENCE OF 83-225 FROM N.A.  
 CC RC STRAIN=C57BL/6J; TISSUE=Retina;  
 CC RX MEDLINE=21085660; PubMed=11217651;  
 CC RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh Y.,  
 CC RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana T.,  
 CC RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 CC RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 CC RA Felschmann W., Gaasterland T., Glass C., Kling B., Kochiwa H.,  
 CC RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 CC RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 CC RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 CC RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 CC RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 CC RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 CC RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 CC RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,  
 CC RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 CC RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
 CC RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 CC RA Hayashizaki Y.;  
 CC RA "Functional annotation of a full-length mouse cDNA collection."  
 CC RL Nature 409:685-690(2001).  
 CC  
 CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/Ap03. Weak  
 CC inducer of apoptosis in some cell types. Promotes angiogenesis and  
 CC the proliferation of endothelial cells. Mediates NF-kappaB  
 CC activation (By similarity).  
 CC -1- SUBUNIT: Homotrimer (Potential).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
 CC similarity).  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- PTM: The soluble form is produced from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF030100; AAC53517.1; -;  
 CC EMBL: AK020909; BAB32249.1; -;  
 CC MGD: MGI:1196259; Tnfsf12.  
 CC DR InterPro: IPR000478; TNF\_family.  
 CC DR Pfam: PF00229; TNF\_1.  
 CC DR SMART: SM00207; TNF\_1.  
 CC DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC DR PROSITE: PS00049; TNF\_2; 1.  
 CC KW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC FT NON TER 1 1  
 CC FT CHAIN 1 225  
 CC FT CHAIN 70 225  
 CC FT CHAIN 70 225  
 CC FT TRANSMEM <1 21  
 CC FT SITE 22 225  
 CC FT SITE 69 70  
 CC  
 CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC MEMBER 12, MEMBRANE FORM.  
 CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC MEMBER 12, SECRETED FORM (BY SIMILARITY).  
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC (POTENTIAL).  
 CC EXTRACELLULAR (POTENTIAL).  
 CC CLEAVAGE (BY SIMILARITY).

FT DISULFID 167 186 POTENTIAL.  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 225 AA; 24781 MW; 90C412CC0480659B CRC64;

Query Match 11.3%; Score 32; DB 1; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-23;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RRAIAHYEHVPRGOGAGAGVGYSGWEE 170  
 |||||||  
 DB 80 RRAIAHYEHVPRGOGAGAGVGYSGWEE 111

## RESULT 3

YFBW\_ECOLI  
 ID YFBW\_ECOLI STANDARD; PRT; 111 AA.  
 AC Q47377;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yfbw.  
 GN YFBW OR B2257.1.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE OF 81-111 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=96186953; PubMed=8626063;  
 RA Sharma V., Hudepeth M.E., Meganathan R.;  
 RT "Menadiolone (vitamin K2) biosynthesis: localization and  
 RT characterization of the menE gene from Escherichia coli.";  
 RL Gene 168:43-48(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE000315; NOT ANNOTATED\_CDS.  
 DR EMBL: L35031; BAB04895.1; -  
 DR Ecogene; EG14344; yfbw.  
 DR InterPro: IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 36 56 POTENTIAL.  
 FT TRANSMEM 62 82 POTENTIAL.  
 FT TRANSMEM 88 108 POTENTIAL.  
 SQ SEQUENCE 111 AA; 12192 MW; 7CEA06D75DA33D69 CRC64;

Query Match 3.5%; Score 10; DB 1; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALACTGL 72  
 |||||||  
 DB 40 LGLALACTGL 49

## RESULT 4

PSAB\_ODOSI  
 ID PSAB\_ODOSI STANDARD; PRT; 733 AA.  
 AC P49480;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Photosystem I P700 chlorophyll A apoprotein A2 (psab) (PST-B).  
 GN PSAB.  
 OS Odontella sinensis (Marine centric diatom).  
 OC Chloroplast.  
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
 OC Biddulphiophycidae; Eupodiscaceae; Eupodiscaceae; Odontella.  
 OX NCBI\_TaxID=2839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;  
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,  
 RT Odontella sinensis.";  
 RL Plant Mol. Biol. Rep. 13:336-342(1995).  
 CC -1- FUNCTION: Psab and psab bind P700, the primary electron donor of  
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and  
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin  
 CC oxidoreductase.  
 CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1  
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.  
 CC -1- SUBUNIT: A psab/B heterodimer binds the P700 chlorophyll special  
 CC pair and subsequent electron acceptors. The PSI reaction center of  
 CC higher plants and algae is composed of one at least 11 subunits.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast  
 CC thylakoid membrane.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL: Z67753; CA91749.1; -  
 DR HSSP: P25897; LJB0.  
 DR InterPro: IPR001280; PSI\_Psaa/B.  
 DR Pfam; PF00223; psaa-psab; 1.  
 DR PRINTS; PR00257; PHOTOSPSAB.  
 DR PROSITE; PS00419; PHOTOSYSTEM\_I\_PsAB; 1.  
 KW Photosynthesis; Photosystem I; Transport; Electron transport;  
 KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
 FT TRANSMEM 46 69 I (POTENTIAL).  
 FT TRANSMEM 134 157 II (POTENTIAL).  
 FT TRANSMEM 174 198 III (POTENTIAL).  
 FT TRANSMEM 272 290 IV (POTENTIAL).  
 FT TRANSMEM 329 352 V (POTENTIAL).  
 FT TRANSMEM 368 394 VI (POTENTIAL).  
 FT TRANSMEM 416 438 VII (POTENTIAL).  
 FT TRANSMEM 516 534 VIII (POTENTIAL).  
 FT TRANSMEM 574 595 IX (POTENTIAL).  
 FT TRANSMEM 642 664 X (POTENTIAL).  
 FT TRANSMEM 706 726 XI (POTENTIAL).  
 FT TRANSMEM 766 786 X (POTENTIAL).  
 FT METAL 558 558 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
 FT METAL 567 567 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
 FT BINDING 653 653 P700 SPECIAL PAIR CHLOROPHYLL AXIAL  
 FT BINDING 653 653 LIGAND (BY SIMILARITY).  
 FT BINDING 661 661 A0 CHLOROPHYLL (BY SIMILARITY).  
 FT BINDING 669 669 A0 CHLOROPHYLL (BY SIMILARITY).  
 FT BINDING 670 670 A1 PHYTYLOQUINONE (BY SIMILARITY).  
 FT BINDING 699 699 A1 PHYTYLOQUINONE (BY SIMILARITY).  
 SQ SEQUENCE 733 AA; 82103 MW; 13439AF1E441BB7 CRC64;

Query Match

3.2%; Score 9; DB 1; Length 733;

Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIACLG 71  
| | | | |  
Db 333 LGIACLG 341

## RESULT 5

PSAB\_CVACA STANDARD; PRT; 734 AA.

ID PSAB\_CVACA  
AC 09106:  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).  
GN PSAB.  
OS Cyanidium caldarium.  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Porphyridiales; Porphyridiaceae;  
OC Cyanidium.  
OX NCBI\_TaxID=2771;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RK-1;  
RX MEDLINE=20496959; PubMed=11040290;  
RA Gloeckner G., Rosenthal A., Valentin K.-U.;  
RT "The structure and gene repertoire of an ancient red algal plastid genome".  
RT J. Mol. Evol. 51:382-390(2000).

CC -1- FUNCTION: PsaA and psab bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.  
CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.  
CC -1- SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.  
CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AF022186; AAF12881.1;  
CC HSSP: P25897; LJB0.  
CC InterPro: IPR001280; PSI\_PsaA/B.  
CC Pfam: PF00223; psaa-psab.1.  
CC PRINTS: PR00257; PHOTOSPSAB.  
CC PROSITE: PS00419; PHOTOSYSTEM\_I\_PSAAB; 1.  
CC Photosynthesis; Photosystem I; Transport; Electron transport;  
KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
KW Iron-sulfur; 4Fe-4S; Chlorophyll.

FT TRANSMEM 46 69 I (POTENTIAL).  
FT TRANSMEM 135 158 II (POTENTIAL).  
FT TRANSMEM 175 199 III (POTENTIAL).  
FT TRANSMEM 273 291 IV (POTENTIAL).  
FT TRANSMEM 330 353 V (POTENTIAL).  
FT TRANSMEM 369 395 VI (POTENTIAL).  
FT TRANSMEM 417 439 VII (POTENTIAL).  
FT TRANSMEM 517 535 VIII (POTENTIAL).  
FT TRANSMEM 575 596 IX (POTENTIAL).  
FT TRANSMEM 643 665 X (POTENTIAL).  
FT TRANSMEM 707 727 XI (POTENTIAL).  
FT METAL 559 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
FT METAL 568 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).

FT BINDING 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL  
FT BINDING 662 662 LIGAND (BY SIMILARITY).  
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).  
FT BINDING 671 671 A0 CHLOROPHYLL (BY SIMILARITY).  
FT BINDING 700 700 A1 PHYTYLOQUINONE (BY SIMILARITY).  
FT BINDING 700 700 A1 PHYTYLOQUINONE (BY SIMILARITY).  
SQ SEQUENCE 734 AA; 82359 MW; 4496AA2AE39CA9B9 CRC64;

Query Match 3.2%; Score 9; DB 1; Length 734;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIACLG 71  
| | | | |  
Db 334 LGIACLG 342

## RESULT 6

ADHS\_GLUOX STANDARD; PRT; 179 AA.

ID ADHS\_GLUOX  
AC 005544;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Alcohol dehydrogenase 15 kDa subunit precursor (G3-ADH subunit III).  
GN ADHS.  
OS Gluconobacter oxydans (Gluconobacter suboxydans).  
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;  
OC Gluconobacter.  
OX NCBI\_TaxID=442;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-40.  
RC STRAIN-IFO 12528;  
RX MEDLINE=97208225; PubMed=9055427;  
RA Kondo K., Horinouchi S.;

RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus".  
RT Appl. Environ. Microbiol. 63:1131-1138(1997).  
CC -1- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.  
CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: D86440; BAA19756.1;  
CC Membrane; Ferriplasmic; Signal.  
CC SIGNAL 1 24  
FT CHAIN 25 179 ALCOHOL DEHYDROGENASE 15 KDA SUBUNIT.  
FT MOD.RES 25 25 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 179 AA; 19943 MW; F6AF24365B3C66 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 179;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LAIGIGLA 66  
| | | | |  
Db 11 LAIGIGLA 18

RESULT 7  
ID Y304\_BROME STANDARD; PRT; 220 AA.  
AC Y304\_BROME  
AC 08YD73;

```

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein BME110304.
GN BME110304.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; Pubmed=11756588;
RA Delvecchio V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonksi L., Larsen N., D'Souza M., Bernal A., Goldsman E.,
RA Salkov E., Blzer P.H., Hagius S., O'Callaghan D., Lelesson J.-J.,
RA Hesselkorn R., Kyriades N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0191 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE009669; AAL53546.1;
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 39
FT TRANSMEM 54 72
FT TRANSMEM 85 104
FT TRANSMEM 124 146
FT TRANSMEM 153 175
FT TRANSMEM 179 198
SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17FA CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 220;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ALVPLAL 61
Db 132 ALVPLAL 139
|||||

```

```

CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ245705; CAB53398.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHDOPSN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT TRANSMEM 1 37
FT TRANSMEM 38 63
FT TRANSMEM 64 72
FT TRANSMEM 73 93
FT TRANSMEM 94 118
FT TRANSMEM 119 140
FT TRANSMEM 141 163
FT TRANSMEM 164 183
FT TRANSMEM 184 191
FT TRANSMEM 192 211
FT TRANSMEM 212 240
FT TRANSMEM 241 266
FT TRANSMEM 267 279
FT TRANSMEM 280 300
FT TRANSMEM 301 317
FT TRANSMEM 317 329
FT TRANSMEM 329 345
FT TRANSMEM 345 365
SQ SEQUENCE 317 AA; 34699 MW; 6615D2146E1D247F CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 317;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 RARRAIA 144
Db 160 RARRAIA 167
|||||

```

```

RESULT 9
ID FDBB_METUA STANDARD; PRT; 379 AA.
AC Q60316;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase beta chain (EC 1.2.1.2).
GN MD0005.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; Pubmed=8688087;
RA Butt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uitterlank T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

```

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RL *jannaschii*.";  
 CC Science 273:1058-1073(1996).  
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.  
 CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLE).  
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).  
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
 CC -1- SIMILARITY: STRONG, TO THE BETA SUBUNIT OF *M.THERMOAUTOTROPHICUM*  
 CC FDH.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U67459; AAB97986.1; -  
 CC DR TIGR: M20005; -  
 CC DR InterPro: IPR001450; 4Fe4S\_ferredoxin.  
 CC Pfam: PF000037; fer4; 1.  
 CC DR PROSITE: PS00198; 4FE4S\_FERREDOXIN, 2.  
 CC KW Hypothetical protein; Oxidoreductase; NAD; Electron transport;  
 CC KM Iron-sulfur; 4Fe-4S; Complete proteome.  
 CC FT METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 283 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 286 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 290 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 330 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 333 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 336 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 340 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC SO SEQUENCE 379 AA; 43014 MW; 9C257CCAD5547F5A CRC64;  
 CC -----  
 CC Query Match 2.8%; Score 8; DB 1; Length 379;  
 CC Best Local Similarity 100.0%; Pred. No. 7.8;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC 217 LVDSVIAL 224  
 CC Db 35 LVDSVIAL 42  
 CC -----  
 CC RESULT 10  
 CC ZP3\_MESAU STANDARD; PRT; 422 AA.  
 CC AC P23491;  
 CC DT 01-NOV-1991 (Rel. 20, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 CC DE zona pellucida sperm-binding protein 3 precursor (zona pellucida  
 CC DE glycoprotein zp3) (Sperm receptor) (zona pellucida protein C).  
 CC GN ZP3.  
 CC OS Mesocricetus auratus (Golden hamster).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC OC Mesocricetus.  
 CC OX NCBI\_TaxID=10036;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX TISSUE-Ovary;  
 CC RX MEDLINE=91078540; PubMed=2257975;  
 CC RA Klenk H.-P., Ruitz-Seller B., Wasserman P.M.;  
 CC RA "Genomic organization and polypeptide primary structure of zona  
 CC RA pellucida glycoprotein hzp3, the hamster sperm receptor.";  
 CC RL Dev. Biol. 142:414-421(1990).  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR  
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE  
 CC SPECIES-SPECIFICITY OF THE INSEMINATION.  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN

CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular  
 CC matrix.  
 CC -1- TISSUE SPECIFICITY: OOCYTES.  
 CC -1- DEVELOPMENTAL STAGE: GROWING OOCYTES.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M63629; AAA37079.1; -  
 CC DR InterPro: IPR001507; Endoglin/CD105.  
 CC Pfam: PF00100; zona-pellucida; 1.  
 CC DR PRINTS: PR00023; ZPELLUCIDA.  
 CC DR SMART: SM00241; ZP; 1.  
 CC DR PROSITE: PS00682; ZP\_DOMAIN.  
 CC KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;  
 CC KM Extracellular matrix.  
 CC FT SIGNAL 1 22  
 CC FT CHAIN 23 422  
 CC FT DOMAIN 23 386  
 CC FT TRANSMEM 387 407  
 CC FT DOMAIN 408 422  
 CC FT DOMAIN 45 306  
 CC FT DOMAIN 119 158  
 CC FT DOMAIN 208 257  
 CC FT CARBOHYD 146 146  
 CC FT CARBOHYD 271 271  
 CC FT CARBOHYD 302 302  
 CC SO SEQUENCE 422 AA; 45827 MW; D0F95B57FE87EB01 CRC64;  
 CC -----  
 CC Query Match 2.8%; Score 8; DB 1; Length 422;  
 CC Best Local Similarity 100.0%; Pred. No. 8.6;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC 59 LALGGLA 66  
 CC Db 386 LALGGLA 393  
 CC -----  
 CC RESULT 11  
 CC CYDC\_HAEIN STANDARD; PRT; 576 AA.  
 CC AC P45081;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Transport ATP-binding protein cydc.  
 CC GN CYDC OR H1156.  
 CC OS Haemophilus influenzae.  
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC OC Haemophilus.  
 CC OX NCBI\_TaxID=727;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX STRAIN=RD / KW20 / ATCC 51907;  
 CC RX MEDLINE=95350630; PubMed=7542800;  
 CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 CC RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 CC RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 CC RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,  
 CC RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 CC RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,  
 CC RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 CC RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 CC RA Venter J.C.;  
 CC "Whole-genome random sequencing and assembly of *Haemophilus influenzae*



```

RT Rd.":
RL Science 269:496-512(1995).
CC -1- FUNCTION: SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC
CC RESPIRATION. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32795; AAC22811.1; -.
DR TIGR: H1156; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR01140; ABCtransportTM.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00664; ABC_membrane; 1.
DR ProDom: PD000006; ABC_transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT NP_BIND 372 379 ATP (POTENTIAL).
SQ SEQUENCE 576 AA; 64831 MW; A9ACDB9B9294B1B3 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 PLALGIGL 65
DB 159 PLALGIGL 166

RESULT 12
GGTS_HUMAN
ID GGT5_HUMAN STANDARD; PRT; 586 AA.
AC P36269; O96PCL;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-glutamyltranspeptidase 5 precursor (EC:2.3.2.2) (Gamma-
DE glutamyltransferase 5) (GGR-Rel).
GN GGT1 OR GGT5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA MEDLINE-91296809; PubMed-1676642;
RA Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,
RA Groffen J.;
RT "Identification of a human gamma-glutamyl cleaving enzyme related to,
RT but distinct from, gamma-glutamyl transpeptidase."
RL Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;

```

```

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CAN HYDROLYZE THE GAMMA-GLUTAMYL MOIETY OF GLUTATHIONE;
CC AS WELL AS CONVERT LEUKOTRIENE C4 TO LEUKOTRIENE D4.
CC -1- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid =
CC peptide + 5-L-glutamyl-amino acid.
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -1- SIMILARITY: STRONG, TO OTHER GGT ENZYMES AND TO GL-7ACA ACYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64099; AA58503.1; -.
DR EMBL: BC011362; AAH11362.1; -.
DR PIR: A41125; A41125.
DR MEROPS: T03.002; -.
DR GeneW: HGNC:4260; GGT1A1.
DR MIM: 137168; -.
DR InterPro: IPR00101; Glutinsptidase.
DR Pfam: PF01019; G-glu_transpept; 1.
DR PRINTS: PR01210; GGTTRANSPASE.
DR PROSITE: PS00462; G-GLU_TRANSPPEPTIDASE; 1.
DR Transfaser: Acyltransferase; Zymogen; Glycoprotein; Transmembrane;
KW Glutathione biosynthesis; Signal-anchor.
FT CHAIN 1 387 GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 HEAVY
FT FT CHAIN (POTENTIAL).
FT FT CHAIN (POTENTIAL).
FT FT CHAIN (POTENTIAL).
FT TRANSMEM 7 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (PROBABLE).
FT DOMAIN 30 586 LDMENAL (POTENTIAL).
FT CARBOHD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 437 437 R -> K (IN REF. 2).
FT CONFLICT 437 437 W -> R (IN REF. 2).
SQ SEQUENCE 586 AA; 62319 MW; 1BE543CB0934B16B CRC64;

Query Match 2.8%; Score 8; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGIGIALA 68
DB 14 LGIGIALA 21

RESULT 13
HSP1_SAGIM
ID HSP1_SAGIM STANDARD; PRT; 49 AA.
AC P24714;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm protamine P1 (Cysteine-rich protamine).
GN PRM1.
OS Saguinus imperator (Tamarin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
OX NCBI_TaxID:9491;
RN [1]
RP SEQUENCE FROM N.A.

```

RC TISSUE-liver;  
 RA MEDLINE-92051332; PubMed-1840669;  
 RX Queralt R., Oliva R.;  
 RT "Protamine 1 gene sequence from the primate *Saguinus imperator*  
 RT isolated with PCR using consensus oligonucleotides.";  
 RL Nucleic Acids Res. 19:5786-5786(1991).  
 CC -1- FUNCTION: PROTAminES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE  
 CC DNA-HELIX (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: Testis.  
 CC -1- SIMILARITY: BELONGS TO THE PROTAminE P1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X61678; CAA3853.1; -;  
 DR PIR: S22582; S22582.  
 DR InterPro: IPR000221; Protamine\_P1.  
 DR Pfam: PF00260; Protamine\_P1; 1.  
 DR PROSITE: PS00048; PROTAminE\_P1; 1.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.  
 FT INIT\_MET 0  
 FT SEQUENCE 49 AA; 6545 MW; 8399C403F5B207F6 CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 42 QRRRGRR 48  
 DB 17 QRRGR 23  
 RESULT 14  
 HSP1\_DIDMA STANDARD; PRT; 57 AA.  
 AC P35305;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sperm protamine P1.  
 GN PPM1.  
 OS *Delphinus marsupialis virginiana* (North American opossum), and  
 OS *Monodelphis domestica* (Short-tailed grey opossum).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Eutelestomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267; 13616;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-D. marsupialis;  
 RX MEDLINE-93345500; PubMed-8344286;  
 RA Winkfein R.J., Nishikawa S., Connor W., Dixon G.H.;  
 RT "Characterization of a marsupial sperm protamine gene and its  
 RT transcripts from the North American opossum (*Didelphis*  
 RT *marsupialis*).";  
 RL Eur. J. Biochem. 215:63-72(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-M. domestica;  
 RX MEDLINE-95215351; PubMed-7700877;  
 RA Redif J.D., Krajewski C., Western M., Winkfein R.J., Dixon G.H.;  
 RT "Molecular phylogeny and evolution of marsupial protamine P1 genes";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 259:7-14(1995).  
 CC -1- FUNCTION: PROTAminES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF

CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: Testis.  
 CC -1- SIMILARITY: BELONGS TO THE PROTAminE P1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L17007; AAA02812.1; -;  
 DR EMBL: X74044; CAA52193.1; -;  
 DR EMBL: L35448; AAA74612.1; -;  
 DR PIR: S34045; S34045.  
 DR InterPro: IPR000221; Protamine\_P1.  
 DR Pfam: PF00260; Protamine\_P1; 1.  
 DR PROSITE: PS00048; PROTAminE\_P1; 1.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.  
 FT INIT\_MET 0  
 FT SEQUENCE 57 AA; 7810 MW; 283715B280214E52 CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 43 RRRRG 49  
 DB 34 RRRGRG 40  
 RESULT 15  
 A62F\_DROME STANDARD; PRT; 115 AA.  
 AC Q46202;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Accessory gland protein Acp62F precursor.  
 GN ACP62F OR CG1262.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN-Canton-S; TISSUE-Male accessory gland;  
 RX MEDLINE-98135120; PubMed-9474779;  
 RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neudam D.M., Park M., Trim V.K.;  
 RT "New genes for male accessory gland proteins in *Drosophila*  
 RT *melanogaster*.";  
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Berkeley;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailes R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwan C.,  
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., Mcherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 Palazzolo M., Pitman G.S., Pan S., Pollard J., Port V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 [3]  
 SEQUENCE OF 7-111 FROM N.A.  
 RC STRAIN-Various strains;  
 RX MEDLINE-20556153; PubMed-11102381;  
 RA Begun D.J., Whitely P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;  
 RT "Molecular population genetics of male accessory gland proteins in  
*Drosophila*.";  
 RL Genetics 156:1879-1888(2000).  
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
 MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL  
 FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO  
 AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE  
 AND ECG RELEASE.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: SEMINAL FLUID.  
 CC -1- SIMILARITY: SOME, TO P. NIGRIVENTER TX2-6.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; 085763; AAB96387.1; -  
 DR EMBL; AE003475; AAF47683.1; -  
 DR EMBL; AY010608; AAG35367.1; -  
 DR EMBL; AY010609; AAG35368.1; -  
 DR EMBL; AY010610; AAG35369.1; -  
 DR EMBL; AY010611; AAG35370.1; -  
 DR EMBL; AY010612; AAG35371.1; -  
 DR EMBL; AY010613; AAG35372.1; -  
 DR EMBL; AY010614; AAG35373.1; -  
 DR EMBL; AY010615; AAG35374.1; -  
 DR EMBL; AY010616; AAG35375.1; -  
 DR EMBL; AY010617; AAG35376.1; -  
 DR FlyBase; FBgn0020509; ACP62F.  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR Pfam; PF01826; TIL\_1.  
 KW Behavior: Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.  
 SO SEQUENCE 115 AA; 12570 MW; 4326AA6FC32291D CRC64;

Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 68 ACGILL 74  
 Db 11 ACGILL 17  
 RESULT 16  
 Y151\_ECOLI STANDARD; PRT; 118 AA.  
 AC P03838;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Insertion element IS5 very hypothetical 12 kDa protein.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82028652; PubMed-6269959;  
 RA Schöner B., Kahn M.;  
 RT "The nucleotide sequence of IS5 from *Escherichia coli*.";  
 RL Gene 14:165-174(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82028652; PubMed-6269958;  
 RA Engler J.A., van Bree M.P.;  
 RT "The nucleotide sequence and protein-coding capability of the  
 transposable element IS5.";  
 RL Gene 14:155-163(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
 RA Davis K., Federspiel N., Hyman R., Kaiman S., Komp C., Kurd O.,  
 RA Laszkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
 RA Davis R.W.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,  
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R.,  
 RA Ichihara S., Mizuno T., Makino K., Nakata A., Yura T., Sempel G.,  
 RA Mischuchi K.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-97251357; PubMed-9097039;  
 RA Alta H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,  
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
 RA Motomura K., Nakano S., Nakamura Y., Nishimoto H., Nishio Y.,  
 RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubram S.,  
 RA Yamamoto Y., Takeda J., Takemoto K., Takeuchi Y., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome  
 corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J01734; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U70214; AAB08680.1; -;

Query Match 2.5%; Score 7; DB 1; Length 115;

DR EMBL: D83536; -: NOT\_ANNOTATED\_CDS.  
 DR EMBL: D90771; BAA14925.1; -  
 DR EMBL: D90772; BAA14935.1; -  
 DR EMBL: D90831; BAA15715.1; -  
 DR EMBL: D90841; BAA15872.1; -  
 DR EMBL: D90847; BAA15958.1; -  
 DR EMBL: D90848; BAA15963.1; -  
 DR PIR: A04466; IECSB.  
 KW Hypothetical protein; Transposable element.  
 SQ SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DCGAVRQ 29  
 1111111  
 Db 107 DCGAVRQ 113

RESULT 17  
 IL13\_MOUSE STANDARD; PRT; 131 AA.  
 AC P20109;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).  
 GN IL13 OR IL-13.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RX MEDLINE=89093958; Pubmed=2521353;  
 RA Brown K.D., Zurewski S.M., Mosmann T.R., Zurewski G.;  
 RT "A family of small inducible proteins secreted by leukocytes are members of a new superfamily that includes leukocyte and fibroblast-derived inflammatory agents, growth factors, and RT indicators of various activation processes."  
 RL J. Immunol. 142:679-687(1989).  
 CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION. SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS. MAY BE CRITICAL IN REGULATING INTERFERON-GAMMA SYNTHESIS. (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: M23504; AAA40149.1; -  
 DR PIR: E30552; E30552.  
 DR HSSP: P35225; 31TR.  
 DR MGD: MGI:96541; 1113.  
 DR InterPro: IPR003634; Interleukin\_13.  
 DR InterPro: IPR001325; Interleukin\_4\_13.  
 DR Pfam: PF03487; Interleukin\_13; 1.  
 DR ProDom: PD015987; Interleukin\_13; 1.  
 DR SMART: SM00190; IL4\_13; 1.  
 DR PROSITE: PS00838; INTERLEUKIN\_4\_13; 1.  
 KW Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 131 INTERLEUKIN-13.  
 FT DISULFID 51 79 BY SIMILARITY.  
 FT DISULFID 67 93 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 131 AA; 14107 MW; 954F93F105713BED CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71  
 1111111  
 Db 9 LALACLG 15

RESULT 18  
 IL13\_RAT STANDARD; PRT; 131 AA.  
 ID IL13\_RAT  
 AC P42203;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).  
 GN IL13 OR IL-13.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney cortex;  
 RX MEDLINE=94092138; Pubmed=7916615;  
 RA Lakkis F.G., Cruet E.N.;  
 RT "Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expression in experimental glomerulonephritis."  
 RL Biochem. Biophys. Res. Commun. 197:612-618(1993).  
 CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION. SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS. MAY BE CRITICAL IN REGULATING INTERFERON-GAMMA SYNTHESIS. (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: L26913; AAA16478.1; -  
 DR HSSP: P35225; 31TR.  
 DR InterPro: IPR003634; Interleukin\_13.  
 DR InterPro: IPR001325; Interleukin\_4\_13.  
 DR Pfam: PF03487; Interleukin\_13; 1.  
 DR ProDom: PD015987; Interleukin\_13; 1.  
 DR SMART: SM00190; IL4\_13; 1.  
 DR PROSITE: PS00838; INTERLEUKIN\_4\_13; 1.  
 KW Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 131 INTERLEUKIN-13.  
 FT DISULFID 52 80 BY SIMILARITY.  
 FT DISULFID 68 94 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 131 AA; 14093 MW; E5008CAB8DE8C201 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71

DB 9 LALACIG 15

RESULT 19  
YK01\_PYRHO STANDARD: PRT; 147 AA.

AC 057781;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PH2001.

OS Pyrococcus horikoshii.  
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.

OK NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=96844137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).

CC EMBL: AP000001; BAA31940.1; -;  
DR EMBL: AP000007; BAA31940.1; JOINED.  
DR EMBL: AP000007; BAA31943.1; -;  
DR EMBL: AP000001; BAA31943.1; JOINED.  
KW Hypothetical protein: Transmembrane: Complete proteome.  
FT TRANSMEM 41 61 POTENTIAL.  
FT TRANSME 67 87 POTENTIAL.  
SQ SEQUENCE 147 AA; 15324 MW; 247ED12FCEFE265B9 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 147;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LALCIGL 73  
DB 45 LALCIGL 51

RESULT 20  
TNFC\_PIG STANDARD: PRT; 150 AA.  
AC 097SV8;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
DE necrosis factor ligand superfamily member 3) (Fragment).  
GN LTB OR TNFSF3 OR TNFC.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteria; Cetartiodactyla; Suidae; Sus.  
OK NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=large white; TISSUE=Fibroblast;  
RX MEDLINE=21108615; PubMed=11169259;

RA Chardon P., Rogel-Galliard C., Cattolico L., Duprat S., Vaiman M.,  
RA Renard C.;  
RT "Sequence of the swine major histocompatibility complex region  
RT containing all non-classical class I genes.";  
RL Tissue Antigens 57:55-65(2001).

CC -1- FUNCTION: Cytokine that binds to LTB8/TNFRSP3. May play a specific  
CC role in immune response regulation. Provides the membrane anchor  
CC for the attachment of the heterotrimeric complex to the cell  
CC surface.

CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
CC (less prevalent) two LTA and one LTB subunits.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).

CC EMBL: AJ251914; CAB63851.1; -;  
DR InterPro: IPR003636; TNF\_abc.  
DR InterPro: IPR000478; TNF\_family.  
DR Pfam: PF00229; TNF; 1.  
DR PRINTS: PR01234; TNECROSISFCT.  
DR PRODOM: PD002012; TNF\_abc; 1.  
DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF\_1; 1.  
DR PROSITE: PS50049; TNF\_2; 1.  
KW Cytokine; Cytotoxin; Glycoprotein.  
FT NON\_TER 1 1  
FT CARBOHYD 128 128 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 150 AA; 16423 MW; FE5C4CC657658B48 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 150;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GXYLYC 200  
DB 37 GLYLYC 43

RESULT 21  
RA05\_ORYSA STANDARD: PRT; 157 AA.  
ID RA05\_ORYSA  
AC 001881;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Seed allergenic protein RA5 precursor.  
GN RA5.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaeae; Oryza.  
OK NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
AC TISSUE=Seed;  
RX MEDLINE=93144699; PubMed=7678765;  
RA Adachi T., Izumi H., Yamada T., Tanaka K., Takeuchi S.,  
RA Nakamura R., Matsuda T.;  
RT "Gene structure and expression of rice seed allergenic proteins  
RT belonging to the alpha-amylase/trypsin inhibitor family.";  
RL Plant Mol. Biol. 21:239-248(1993).  
CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPsin/ALPHA-AMYLASE INHIBITOR  
CC FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11430; BAA01996.1; -
DR PIR: S31078; S31078.
DR HSSP: P01085; 1HSS.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; try/amy1_inhbt.
DR Pfam: PF00234; try_alpha_amy1; 1.
DR PRINTS: PR00808; AMLASEINHBT.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
DR Allergen: MultiGene family; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 157 SEED ALLERGENIC PROTEIN RAS.
SQ SEQUENCE 157 AA; 17118 MW; C8A5495FBF8399E6 CRC64;

Query Match 2.58; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78
Db 11 LLLAVS 17

RESULT 22
HLPA_ECOLI STANDARD; PRT; 161 AA.
ID HLPA_ECOLI
AC P11457;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Histone-like protein HLP-1 precursor (DNA-binding 17 kDa protein).
GN HLPA OR SKP OR OMPH OR B0178 OR Z0190 OR ECS0180.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.
RX MEDLINE=86329735; PubMed=2843433;
RA Holck A., Kleppe K.;
RT "Cloning and sequencing of the gene for the DNA-binding 17K protein
RT of Escherichia coli."
RL Gene 67:117-124(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden W.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Samped G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 5.0 min (189,987 - 281,416bp) region."
RT Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RL [4]

```

```

RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federpiel N., Hyman R., Kalman S., Komp C., Kurl O.,
RA Iashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RL [6]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11288796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RT DNA Res. 8:11-22(2001).
RL [7]
RP SEQUENCE OF 72-161 FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=91100302; PubMed=1987124;
RA Dicker I.B., Seetharam S.R.;
RT "Cloning and nucleotide sequence of the fira gene and the fira200(Ts)
RT allele from Escherichia coli."
RL J. Bacteriol. 173:334-344(1991).
RN [8]
RP SEQUENCE OF 21-32.
RC STRAIN-K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997).
RN [9]
RP SIMILARITY TO S.TYRPHIMRIUM OMPH.
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
RT constituent".
RL FEBS Lett. 262:123-126(1990).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
CC THE OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M21118; AAA24630.1; -
DR EMBL: AE000127; AAC73289.1; -
DR EMBL: D83536; BAA77853.1; -
DR EMBL: U70214; AAB08607.1; -
DR EMBL: AE005194; AAG54460.1; -
DR EMBL: AF002550; BAB33603.1; -
DR EMBL: X54797; CAA38567.1; -
DR EMBL: X75465; CAA53207.1; -

```

DR PIR: J70304; DNEC17.  
 DR PIR: S13728; S13728.  
 DR Ecogene: EG10455; h1p.  
 KM DNA-binding; Outer membrane; Signal; Complete proteome.  
 FT SIGNAL 1 20  
 FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-1.  
 SQ SEQUENCE 161 AA; 17688 MW; 2A96BBD83F3E675 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 DB 9 GIGLALA 15

RESULT 23  
 YIM4\_CAEEL STANDARD; PRT; 170 AA.  
 AC P34378;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein D2007.4 in chromosome III.  
 GN D2007.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodertinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Cratton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnson L., Jones M., Kershaw J., Kirsten J., Lalster N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Koopra A., Saunders D., Shownkeen R.,  
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
 RA Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: L16560; AAA27999.1;  
 DR PIR: S44789; S44789.  
 DR WormPep: D2007.4; CE00129.  
 KM Hypothetical protein.  
 SQ SEQUENCE 170 AA; 19396 MW; 22301D7C65638135 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 VIALRCL 227  
 DB 94 VIALRCL 100

RESULT 24

LACB\_BUBBU STANDARD; PRT; 180 AA.  
 ID LACB\_BUBBU  
 AC P02755; O62822;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-lactoglobulin precursor (Beta-Lg).  
 GN LGB.  
 OS Bubalus bubalis (Domestic water buffalo).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bubalus.  
 OX NCBI\_TaxID=89462;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=99304500; PubMed=10376212;  
 RA Das P., Jain S., Nayak S., Apparao K.B.C., Toley S.M., Garg L.C.;  
 RT "Molecular cloning and sequence analysis of the cDNA encoding  
 RT beta-lactoglobulin in Bubalus bubalis."  
 RL DNA Seq. 10:105-108(1999).  
 RN [2]  
 RP SEQUENCE OF 19-180.  
 RA Kolde H.-J., Liberatori J., Braunitzer G.;  
 RT "The amino acid sequence of the water buffalo beta-lactoglobulin."  
 RL Milchwissenschaft 36:83-86(1981).

CC -1- FUNCTION: Primary component of whey, it binds retinol and is  
 CC probably involved in the transport of that molecule.  
 CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS  
 CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted  
 CC in milk.  
 CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AF005429; CA06532.1;  
 DR PIR: A03219; LGB01.  
 DR HSSP: P02754; LB50.  
 DR InterPro: IPR002345; Lipocalin.  
 DR Pfam: PF00061; Lipocalin\_1.  
 DR PRINTS: PR00179; LIPOCALIN.  
 DR PROSITE: PS00213; LIPOCALIN; 1.  
 KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 180 BETA-LACTOGLOBULIN.  
 FT DISULFID 84 178  
 FT DISULFID 124 137  
 FT DISULFID 124 139  
 SQ SEQUENCE 180 AA; 20023 MW; 6836C97B2C2E33CF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIALAC 69  
 DB 8 LGIALAC 14

RESULT 25  
 LACB\_CAPHI STANDARD; PRT; 180 AA.  
 ID LACB\_CAPHI  
 AC P02756;



DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-lactoglobulin precursor (Beta-Ig).  
 GN LGB.  
 OS Capra hircus (Goat).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Capra.  
 OK NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPD. aegagrus; TISSUE=Mammary gland;  
 RX MEDLINE=94042559; PubMed=8226387;  
 RA Folch J.M., Coll A., Sanchez A.;  
 RT "Cloning and sequencing of the cDNA encoding goat beta-lactoglobulin.";  
 RL J. Anim. Sci. 71:2832-2832(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=95213451; PubMed=7699130;  
 RA Folch J., Coll A., Sanchez A.;  
 RT "Complete sequence of the caprine beta-lactoglobulin gene.";  
 RL J. Dairy Sci. 77:3493-3497(1994).  
 RN [4]  
 RP SEQUENCE OF 19-180.  
 RX MEDLINE=80070611; PubMed=511095;  
 RA Preaux G., Braunitzer G., Schrank B., Strangl A.;  
 RT "The amino acid sequence of goat beta-lactoglobulin.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1595-1604(1979).  
 CC -1- FUNCTION: Primary component of whey, it binds retinol and is probably involved in the transport of that molecule.  
 CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted in milk.  
 CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X58471; CAA41385.1; -  
 DR EMBL: Z19569; CAA79623.1; -  
 DR EMBL: Z19570; CAA79624.1; -  
 DR EMBL: Z33881; CAA83946.1; -  
 DR PIR: A03320; LGCT.  
 DR PIR: S14507; S14507.  
 DR PIR: S42800; S42800.  
 DR HSP: P02754; IBOO.  
 DR Interpro: IPR002345; LIPOCALIN.  
 DR Interpro: IPR000566; LIPOCALIN\_CYFABP.  
 DR Pfam: PF00061; LIPOCALIN.1.  
 DR PRINTS: PR00179; LIPOCALIN.  
 DR PROSITE: PS00213; LIPOCALIN.1.  
 KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.  
 FT CHAIN 1 180  
 FT SIGNAL 1 18  
 FT DISULFID 84 178 BETA-LACTOGLOBULIN.  
 FT DISULFID 124 137  
 FT DISULFID 124 139  
 SEQUENCE 180 AA; 19975 MW; C2449BB02A1A80F1 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 63 LGALAC 69  
 DB 8 LGALAC 14  
 RESULT 26  
 LACB-SHEEP  
 ID LACB-SHEEP STANDARD; PRT; 180 AA.  
 AC P02757;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.  
 OS Ovis aries (Sheep), and  
 OS Ovis orientalis musimon (Mouflon).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Ovis.  
 OK NCBI\_TaxID=9940, 9938;  
 RN [1]  
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=88172489; PubMed=3351935;  
 RA Ali S., Clark A.J.;  
 RT "Characterization of the gene encoding ovine beta-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins.";  
 RL J. Mol. Biol. 199:415-426(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A. (BLG 1).  
 RC SPECIES=Sheep;  
 RX MEDLINE=8704827; PubMed=3096387;  
 RA Gaye P., Hue-Delehante D., Mercier J.-C., Soulier S., Vilotte J.-L., Furet J.-P.;  
 RT "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland.";  
 RL Biochimie 68:1097-1107(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A. (BLG 1).  
 RC SPECIES=Sheep;  
 RX MEDLINE=89057492; PubMed=3194215;  
 RA Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;  
 RT "Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene.";  
 RL Nucleic Acids Res. 16:10379-10380(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=91007276; PubMed=1976573;  
 RA Ali S., McElenaghan M., Simons J.P., Clark A.J.;  
 RT "Characterisation of the alleles encoding ovine beta-lactoglobulins A and B.";  
 RL Gene 91:201-207(1990).  
 RN [5]  
 RP SEQUENCE OF 19-180 (BLG 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=80219294; PubMed=6155855;  
 RA Preaux G., Braunitzer G., Kolde H.-J.;  
 RT "Primary structure of ovine beta-lactoglobulin.";  
 RL Arch. Int. Physiol. Biochim. 88:B45-B46(1980).  
 RN [6]  
 RP SEQUENCE OF 19-180 (BLG 3).  
 RC SPECIES=Sheep;  
 RX MEDLINE=89374823; PubMed=2775495;  
 RA Erhardt G., Godovac-Zimmermann J., Conti A.;  
 RT "Isolation and complete primary sequence of a new ovine wild-type beta-lactoglobulin C.";  
 RL Biol. Chem. Hoppe-Seyler 370:757-762(1989).



```

RN [7]
RP SEQUENCE OF 19-180 (B1G B).
RC SPECIES=O. musimom; PubMed=3426802;
RX MEDLINE=88106996; PubMed=3426802;
RA Godovac-Zimmermann J., Conti A., Napolitano L.,
RT "The complete amino-acid sequence of dimeric beta-lactoglobulin from
RL mouflon (Ovis ammon musimon) milk."
CC Biol. Chem. Hoppe-Seyler 368:1313-1319(1987).
CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
CC THAT MOLECULE.
CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X04520; CAA28204.1; -
DR EMBL: X12817; CAA31305.1; -
DR EMBL: X07004; CAA30059.1; ALT. SEQ.
DR EMBL: X07005; CAA30059.1; JOINED.
DR EMBL: X07006; CAA30059.1; JOINED.
DR EMBL: X07007; CAA30059.1; JOINED.
DR EMBL: X07008; CAA30059.1; JOINED.
DR EMBL: X07009; CAA30059.1; JOINED.
DR EMBL: M32236; AAA31510.1; -
DR EMBL: M32232; AAA31510.1; JOINED.
DR EMBL: M32233; AAA31510.1; JOINED.
DR EMBL: M32235; AAA31510.1; JOINED.
DR EMBL: A03221; LGSB.
DR PIR: S00132; S00132.
DR PIR: A30011; A30011.
DR PIR: B30011; B30011.
DR PIR: A25136; A25136.
DR PIR: J00748; J00748.
DR PIR: J00749; J00749.
DR PIR: S02136; S02136.
DR PIR: S04955; S04955.
DR HSSP: P02754; 1BSO.
DR InterPro: IPR002345; Lipocalin.
DR InterPro: IPR000566; Lipocalin_cyfeABP.
DR Pfam: PF00061; Lipocalin; 1.
DR PRINTS: PR00179; LIPOCALIN.
DR PROSITE: PS00213; LIPOCALIN; 1.
KW Milk; Whey; Retinol-binding; Transport; Signal; Lipocalin.
FT SIGNAL 1 18
FT CHAIN 19 180 BETA-LACTOGLOBULIN.
FT DISULFID 84 178
FT DISULFID 124 137
FT DISULFID 124 139 ALTERNATE.
FT VARIANT 38 38 H -> Y (IN LACTOGLOBULIN 2-A AND 3-C).
FT VARIANT 166 166 R -> Q (IN LACTOGLOBULIN 3-C).
SQ SEQUENCE 180 AA; 19921 MW; BABC82E89E757333 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGALAC 69
DB 8 LGALAC 14

RESULT 27
RNFE_SALTY STANDARD; PRT; 230 AA.
ID RNFE_SALTY

```

```

AC O8XEX9;
DE 15-JUN-2002 (Rel. 41, Created)
DE 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNFE OR STM1454 OR STY1668.
OS Salmonella typhimurium, and
OC Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium; STRAIN=LT2 / SCSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Speeth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RT Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jajels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skellon J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RT Nature 413:848-852(2001).
RN [1]
RP Nature 413:852-856(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE008763; AAL20376.1; -
DR EMBL: AL627271; CAD01913.1; -
DR Stycene: SG27272; rnfE.
DR InterPro: IPR003667; Rnf_Mgr.
DR Pfam: PF02508; Rnf_Mgr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 34 56 POTENTIAL.
FT TRANSMEM 69 87 POTENTIAL.
FT TRANSMEM 91 113 POTENTIAL.
FT TRANSMEM 126 148 POTENTIAL.
FT TRANSMEM 183 205 POTENTIAL.
SQ SEQUENCE 230 AA; 24318 MW; E198B4CEA13F249E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLCIA 66
DB 60 ALGLCIA 66

```

Db 38 ALGGLA 44

RESULT 28

RNFE\_ECO57 STANDARD; PRT; 231 AA.

ID RNFE\_ECO57 STANDARD; PRT; 231 AA.

AC P58344;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Electron transport complex protein rnfe.

GN RNFE OR 22642 OR ECS2341.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G., Iii, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grodeck E.V., Davis N.W., Lim A., Dimantanta E.T., Potamouls K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature 409:529-533(2001).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takemi H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12." DNA Res. 8:11-22(2001).

RL [1]

CC -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc, rnfd, rnfe and rnfg (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE003386; AAG56621.1; -

DR EMBL; AP002558; BAB35764.1; -

DR InterPro; IPR003667; Rnf\_Ngr.

DR Pfam; PF02508; Rnf-Ngr; 1.

DR KEGG; Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT DOMAIN 1 38 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 39 59 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 60 62 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 63 83 POTENTIAL.

FT TRANSSEM 84 85 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 86 106 POTENTIAL.

FT TRANSSEM 107 124 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 125 145 POTENTIAL.

FT TRANSSEM 146 181 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 182 202 POTENTIAL.

FT TRANSSEM 203 231 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 231 231 CYTOPLASMIC (POTENTIAL).

SO SEQUENCE 231 AA; 24489 MW; DAA2CA2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;

Best Local Similarity 100.0%; Pred. No. 47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66

Db 38 ALGGLA 44

RESULT 29

RNFE\_ECOLI STANDARD; PRT; 231 AA.

ID RNFE\_ECOLI STANDARD; PRT; 231 AA.

AC P77179;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Electron transport complex protein rnfe.

GN RNFE OR B1632.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G., Iii, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12." Science 277:1453-1474(1997).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-K12;

RX MEDLINE=97251357; PubMed=9097039;

RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N., Sampel G., Seki Y., Sivasubraman S., Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 mln region on the linkage map." DNA Res. 3:363-377(1996).

RL [3]

RN TOPOLOGY.

RP STRAIN-K12 / JM109;

RX MEDLINE=99342054; PubMed=10411911;

RA Seeef A., Johansson M., Wallin E., von Heijne G.;

RT "Divergent evolution of membrane protein topology: the Escherichia coli rnfa and rnfe homologues." Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).

RL [1]

CC -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc, rnfd, rnfe and rnfg (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE000258; AAC74704.1; -

DR EMBL; D90806; BAA15386.1; -

DR EMBL; D90807; BAA15393.1; -

DR EMBL: D90808; BAA15416.1; -  
 DR ECGene: EG13938; rnfE.  
 DR InterPro: IPR003667; Rnf\_Ngr.  
 DR Pfam: PF02508; Rnf\_Ngr; 1.  
 DR Electron transport; Transmembrane; Inner membrane; Complete proteome.  
 KW DOMAIN 1 38 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 39 59 POTENTIAL.  
 FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 63 83 POTENTIAL.  
 FT DOMAIN 84 85 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 86 106 POTENTIAL.  
 FT DOMAIN 107 124 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 125 145 POTENTIAL.  
 FT DOMAIN 146 181 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 182 202 POTENTIAL.  
 FT DOMAIN 203 231 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 231 AA; 24459 MW; CFA37A2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66  
 |||||  
 DB 38 ALGGLA 44

## RESULT 30

RNC\_CONXBU STANDARD: PRT; 233 AA.

AC P51837; 01-OCT-1996 (Rel. 34, created)

DT 01-OCT-1996 (Rel. 34, last sequence update)

DT 15-OCT-2001 (Rel. 40, last annotation update)

DE Ribonuclease III (EC 3.1.26.3) (Rnase III).

GN RNC.

OS Coccidia burnetii.

OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;

CC Coccidia group; Coccidia.

OX NCBI\_Taxid=777;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CBM1C7;

RX MEDLINE=95131751; PubMed=7830573;

RA Zuber M., Hoover T.A., Powell B.S., Court D.L.;

RL "Analysis of the rnc locus of Coccidia burnetii.";

CC MOJ. Microbiol. 14:291-300(1994).

CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING

CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAs (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-

CC phosphonoeester.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: CONTAINS 1 Rnase III DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 Rnase III DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: L27436; AAA69690.1; -

DR InterPro: IPR001159; DS\_RBD.

DR InterPro: IPR000999; Rnase\_3.

DR Pfam: PF00035; dsrm; 1.

DR Pfam: PF00636; Ribonuclease\_3; 1.

DR SMART: SM00358; DSRM; 1.

DR SMART: SM00335; RIBOC; 1.

DR PROSITE: PS50137; DS\_RBD; 1.

DR PROSITE: PS00517; Rnase\_3\_1; 1.

DR PROSITE: PS0142; Rnase\_3\_2; 1.

KW Hydrolase; Nuclease; Endonuclease; RNA-binding.  
 FT DOMAIN 4 126 Rnase III.  
 FT TRANSMEM 204 220 DBM.  
 SQ SEQUENCE 233 AA; 26229 MW; 1A11CB4AD96ED2FF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ARRLPL 16  
 |||||  
 DB 164 ARRLPL 170

## RESULT 31

RNFE\_YERPE STANDARD: PRT; 233 AA.

AC Q8ZED4; 15-JUN-2002 (Rel. 41, created)

DT 15-JUN-2002 (Rel. 41, last sequence update)

DT 15-JUN-2002 (Rel. 41, last annotation update)

DE Electron transport complex protein rnfE.

GN Rnfe OR YP02240.

OS Yersinia pestis.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

CC Yersinia.

OX NCBI\_Taxid=632;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CO-92 / Biovar Orientalis;

RX MEDLINE=21470413; PubMed=11586360;

RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Feltham T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,

RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;

RL "Genome sequence of Yersinia pestis, the causative agent of plague.";

CC Nature 413:523-527(2001).

CC -1- FUNCTION: May be part of a membrane complex involved in electron

CC transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,

CC rnfD, rnfE and rnfG (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: AJ414151; CAC91046.1; -

DR InterPro: IPR003667; Rnf\_Ngr.

DR Pfam: PF02508; Rnf\_Ngr; 1.

DR Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 34 56 POTENTIAL.

FT TRANSMEM 69 87 POTENTIAL.

FT TRANSMEM 91 113 POTENTIAL.

FT TRANSMEM 126 148 POTENTIAL.

FT TRANSMEM 184 206 POTENTIAL.

SQ SEQUENCE 233 AA; 24587 MW; 491E1BF35E8CB90 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;

Best Local Similarity 100.0%; Pred. No. 47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66  
 |||||

```

Db      38 ALGIGLA 44

RESULT 32
RNF_E_HAEN  STANDARD; PRT: 235 AA.
AC  Q57020: P96346;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Electron transport complex protein rnfE.
GN  RNF_E OR H11688.
OS  Haemophilus influenzae.
OC  Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC  Haemophilus.
OX  NCBI_Taxid=727;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-Rd / RM20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Uleberg T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fritchman J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT  Rd."
RL  Science 269:496-512(1995).
CC  -1- FUNCTION: May be part of a membrane complex involved in electron
CC  transport (By similarity).
CC  -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC  rnfD, rnfE and rnfG (By similarity).
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC  (Potential).
CC  -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: U32841; AAC23334.1; -
DR  TIGR: H11688; -
DR  InterPro: IPR003667; Rnf_Nqr.
DR  Pfam: PF02508; Rnf_Nqr; 1.
KW  Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT  TRANSMEM 63 83 POTENTIAL.
FT  TRANSMEM 93 113 POTENTIAL.
FT  TRANSMEM 117 137 POTENTIAL.
FT  TRANSMEM 152 172 POTENTIAL.
FT  TRANSMEM 206 226 POTENTIAL.
SQ  SEQUENCE 235 AA; 25845 MW; C054FE59647837A CRC64;

Query Match 2.58; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      60 ALGIGLA 66
Db      62 ALGIGLA 68

RESULT 33
TN14_MOUSE  STANDARD; PRT: 239 AA.
AC  Q90YH9;

DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Tumor necrosis factor ligand superfamily member 14.
GN  TNFSF14 OR LIGHT.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_Taxid=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  PubMed=10700230;
RX  Tamada K., Shinozaki K., Chapoval A.I., Zhu G., Sica G., Flies D.,
RA  Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;
RT  "Modulation of T-cell-mediated immunity in tumor and graft-versus-host
RT  disease models through the LIGHT co-stimulatory pathway.";
RL  Nat. Med. 6:283-289(2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE-Petal liver;
RA  Mitsuwa K., Nosaka T., Kitamura T., Kojima T.;
RT  "Murine LIGHT, a homologue of human LIGHT which is a member of TNF
RT  family.";
RL  Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE-Lymphoma;
RA  Force W.R., Todd P.K., Mikayama T.;
RT  "Mouse LIGHT: molecular genetics, ligand binding and expression.";
RL  Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the
CC  decoy receptor TNFRSF6B modulates its effects. Activates NFkB
CC  and stimulates the proliferation of T cells.
CC  -1- SUBUNIT: Homotrimer (By similarity).
CC  -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC  similarity).
CC  -1- PTM: The soluble form derives from the membrane form by
CC  proteolytic processing.
CC  -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF123385; AAF76453.1; -
DR  EMBL: AB029155; BAA88559.1; -
DR  EMBL: AF227533; AAF36722.1; -
DR  HSSP: P01375; 4TSV.
DR  MGd: MG11355317; Tnfsl4.
DR  InterPro: IPR003636; TNF_abC.
DR  InterPro: IPR000478; TNF_family.
DR  Pfam: PF00229; TNF; 1.
DR  PRINTS: PR01234; TNECROSISFCR.
DR  PRODOM: PD002012; TNF_abC; 1.
DR  SMART: SM00207; TNF; 1.
DR  PROSITE: PS00251; TNF; 1; FALSE-NEG.
DR  PROSITE: PS50049; TNF_2; 1.
KW  Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.
FT  CHAIN 1 239
FT  CHAIN 782 239
FT  DOMAIN 1 37
FT  TRANSMEM 38 58
FT  TRANSMEM 59 239
FT  SITE 81 82
FT  DISULFID 152 187
FT  CARBOHYD 100 100

```

FT CARBOHYD 191 191 N-LINKED (GIC/NAC... ) (POTENTIAL).  
SQ SEQUENCE 239 AA; 26338 MW; 217874AC71ADBE3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 NRYVPR 127  
Db 221 NRYVPR 227

# RESULT 34

RNFE\_PSEAE STANDARD; PRT: 240 AA.

AC Q9H1B5;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Electron transport complex protein rnfe.  
GN RNFE OR PA3494.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RA MEDLINE=20437337; PubMed=10984043;  
RA Stover C.R., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagron M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.W.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reiter J., Siler M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).

CC -1- FUNCTION: May be part of a membrane complex involved in electron  
transport (By similarity).  
CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc,  
CC rnfd, rnfe and rnfg (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AE004770; AAC06882.1;  
DR InterPro: IPR003667; Rnf Ngr.

DR Pfam: PF02508; Rnf Ngr; 1.  
DR Electron transport; Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 41 61 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 95 115 POTENTIAL.  
FT TRANSMEM 130 150 POTENTIAL.  
FT TRANSMEM 184 204 POTENTIAL.  
SQ SEQUENCE 240 AA; 25794 MW; 3D90687ED462D8B2 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALG1GLA 66  
Db 40 ALG1GLA 46

# RESULT 35

MOEB\_HAEIN STANDARD; PRT: 243 AA.

AC P45211;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Molybdopter in biosynthesis protein moeb.  
GN MOEB OR CHLN OR H1449.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.

OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Rd / KW20 / ATCC 51907;  
RA MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RL Rd.";  
RL Science 269:496-512(1995).

CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COPROCTOR  
(MOLYBDOPTERIN). NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE  
CC IN ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING  
CC FACTOR (MOAF) (BY SIMILARITY).  
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: U32823; AAC3099.1;  
DR TIGR: H11449;  
DR InterPro: IPR00205; NAD binding.  
DR InterPro: IPR000594; THIF-domain.  
DR Pfam: PF00899; THIF; 1.  
KW Molybdenum cofactor biosynthesis; Complete proteome.

SQ SEQUENCE 243 AA; 26996 MW; 218A3382A975BDBD CRC64;

Query Match 2.5%; Score 7; DB 1; Length 243;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185  
Db 10 LRYNRQI 16

# RESULT 36

TNFC\_HUMAN STANDARD; PRT: 244 AA.

AC Q0643; P78370; Q99761;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
DE necrosis factor ligand superfamily member 3).  
GN LTb OR TNFSF3 OR TNFC.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
 RC TISSUE="T-cell";  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brien-Greco B., Foley S.F., Ware C.F.,  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.",  
 RL Cell 72:847-856(1993).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=97445965; PubMed=9299492;  
 RA Marzocha K., Renard N., Charlot C., Blenvenu J., Collier B.,  
 RA Salles G.;  
 RT "Identification of two lymphotoxin beta isoforms expressed in human  
 RT lymphoid cell lines and non-Hodgkin's lymphomas.",  
 RL Biochem. Biophys. Res. Commun. 238:273-276(1997).  
 [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Neville M.J., Milner C.M., Campbell R.D.;  
 RT "A new member of the immunoglobulin superfamily and a V-ATPase G  
 RT subunit are amongst the predicted products of novel genes close to the  
 RT TNF locus in the human MHC.",  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Rowen L., Madan A., Olin S., Shaffer T., James R., Ratcliffe A.,  
 RA Abbas N., Dickhoff R., Loretz C., Madan A., Dots M., Young J.,  
 RA Laszky S., Hood L.;  
 RT "Sequence of the human major histocompatibility complex class III  
 RT region.",  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Shilina S., Tamlya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.",  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLD-70 AND  
 RP PRO-111.  
 RA Rieder M.J., Amel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to LTRB/TNFRSF. May play a specific  
 CC role in immune response regulation. Provides the membrane anchor  
 CC for the attachment of the heterotrimeric complex to the cell  
 CC surface. Isoform 2 is probably non-functional.  
 CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
 CC (less prevalent) one LTB and two LTA subunits.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2, are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: SPLEEN AND THYMUS.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: L11016; AAA9988.1; -  
 DR EMBL: 0689922; AAC51769.1; -  
 DR EMBL: 079029; AAB37342.1; -  
 DR EMBL: L11015; AAA36191.1; -  
 DR EMBL: Y14768; CAA75069.1; -  
 DR EMBL: AF129756; AAD18069.1; -  
 DR EMBL: AF000505; BAB63395.1; -

DR EMBL: AY070219; AAL49954.1; -  
 DR EMBL: AY070219; AAL49955.1; -  
 DR PIR: A46066; A46066.  
 DR HSSP: P01374; 1TNR.  
 DR Genew: HGNC:6711; LTB.  
 DR MIM: 600978; -  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
 KW Alternative splicing; Polymorphism.  
 FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 19 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 49 222 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 222 222 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARSPPLIC 53 77 GLYETADGAGAGAGGIGOKLPEE -> GLGFRSQRRSSO  
 FT VARSPPLIC 78 244 KOISARGSOLPIS (IN ISOFORM 2).  
 FT VARIANT 70 70 MISSING (IN ISOFORM 2).  
 FT VARIANT 70 70 G -> E.  
 FT VARIANT 111 111 /FTID-VAR\_013025.  
 FT VARIANT 111 111 A -> P.  
 FT CONFLICT 60 69 /FTID-VAR\_013026.  
 FT CONFLICT 60 69 DPGAQAGGL -> GLSAGSGGRT (IN REF. 2;  
 FT SEQUENCE 244 AA; 25390 MW; F41569459830ED4C CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 194 GLTYLYC 200  
 DB 131 GLTYLYC 137  
 RESULT 37  
 MOEB\_ECOLI  
 ID MOEB\_ECOLI STANDARD; PRT; 249 AA.  
 AC P12282;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Molycoperoxin biosynthesis protein moeb.  
 GN MOEB OR CHLN OR B0826.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88314906; PubMed=3045084;  
 RA Nohno T., Kasai Y., Saito T.;  
 RT "Cloning and sequencing of the Escherichia coli chln operon involved  
 RT in molycoperoxin biosynthesis.",  
 RL J. Bacteriol. 170:4097-4102(1988).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC SFRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RT Mau B., Shao Y.;  
 RL "The complete genome sequence of Escherichia coli K-12.",  
 RL Science 277:123-1453-1474(1997).  
 [3]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN-K12;
RX Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.,
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC (MOLYBDOPROTEIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M2151; AAA23580.1; -
DR EMBL: AE000185; AAC73913.1; -
DR EMBL: D90720; BAA35514.1; -
DR EMBL: D90721; BAA35521.1; -
DR PIR: B32352; B32352.
DR Ecogene: EG10154; moeb.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; THIF_domain.
DR Pfam: PF00899; Thif; 1.
SQ Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26719 MW; 12C77082B3F9D7D CRC64;

Query Match 2.5%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
DB 10 LRYNRQI 16

RESULT 38
MOEB_SALTY
AC 056057; STANDARD; PRT; 249 AA.
ID 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moeb.
GN MOEB OR STM0845.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602; [1]
RN NCBI STRAIN=LT2;
RP SEQUENCE FROM N.A.
RA Wong K.K., Kwan H.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

```

```

RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC (MOLYBDOPROTEIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U53178; AAA96530.1; -
DR EMBL: AE008735; AAL19781.1; -
DR StyGene: SG1065; moeb.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; THIF_domain.
DR Pfam: PF00899; Thif; 1.
SQ Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26903 MW; 0F0050831D537AD2 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
DB 10 LRYNRQI 16

RESULT 39
KDPX_SERMA
ID KDPX_SERMA STANDARD; PRT; 257 AA.
AC Q54435;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Lipopolysaccharide core biosynthesis glycosyl transferase kdtX
DE (EC 2.---.-).
GN KDPX.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615; [1]
RN NCBI STRAIN=N28B;
RP SEQUENCE FROM N.A.
RX MEDLINE=96422003; PubMed=8824620;
RA Guasch J.F., Pique N., Climent N., Ferrer S., Merino S., Rubires X.,
RA Tomas J.M., Regue M.;
RT "Cloning and characterization of two Serratia marcescens genes
RT involved in core lipopolysaccharide biosynthesis.";
RL J. Bacteriol. 178:5741-5747(1996).
CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WAAE/KDTX
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U52844; AAC44433.1; -  
 DR InterPro: IPR001173; Glycosyltransferase.  
 DR Pfam: PF00535; Glycosyltransferase; Transferrase; Glycosyltransferase.  
 KW Lipopolysaccharide biosynthesis; Transferrase; Glycosyltransferase.  
 SQ SEQUENCE 257 AA; 29233 MW; D40DB57E002FE990 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 257;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 GLLAIV 77  
 |||||  
 Db 229 GLLAIV 235

RESULT 40  
 TNFC\_MARMO STANDARD; PRT; 310 AA.  
 ID TNFC\_MARMO  
 AC O3JML0; O3JML1; 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
 DE necrosis factor ligand superfamily member 3).  
 GN LTB OR TNFSF3 OR TNFC.  
 OS Marmota monax (Woodchuck).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;  
 OC Marmota.  
 OX NCBI\_TaxID=9995;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20184748; PubMed-10721723;  
 RT L.D.H., Havelle E.A., Brown C.L., Cullen J.M.;  
 RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:  
 RT structure, characterization and biological activity";  
 RL Gene 242:295-305(2000).  
 CC -1- FUNCTION: Cytokine that binds to LTBFR/TNFSF3. May play a specific  
 CC role in immune response regulation. Provides the membrane anchor  
 CC for the attachment of the heterotrimeric complex to the cell  
 CC surface.  
 CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
 CC (less prevalent) two LTA and one LTB subunits (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF096268; AAF34866.1; -  
 DR EMBL: AF095587; AAF34865.1; -  
 DR HSSP: P01374; 1NR.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR00478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PSS0049; TNF\_2; 1.  
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.  
 FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 49 310 EXTRACELLULAR (POTENTIAL).

FT CARBOHYD 272 272 N-LINKED (GLCNAC... (POTENTIAL).  
 FT CONFLICT 280 280 D -> H (IN REF. 1; AAF34865).  
 SQ SEQUENCE 310 AA; 32644 MW; 73B354EFC8B3B3BE CRC64;

Query Match 2.5%; Score 7; DB 1; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GLYLVC 200  
 |||||  
 Db 181 GLYLVC 187

RESULT 41  
 ISPH\_XYLFA STANDARD; PRT; 316 AA.  
 ID ISPH\_XYLFA  
 AC O9PAS9;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE ISPH protein.  
 GN ISPH OR LTB OR XP2416.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-9a5c;  
 RX MEDLINE-20365717; PubMed-10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordoli S., Bove J.M., Brites M.R.S.,  
 RA Bueno M.R.P., Camargo A.E., Camargo L.E.A., Carro D.M., Carier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins A.C.Y., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Resquero J.B.,  
 RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsunako M.H.,  
 RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Seubert J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RL Nature 406:151-159(2000).  
 CC -1- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate  
 CC into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate  
 CC (DMAPP) (By similarity).  
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh  
 CC (last) step.  
 CC -1- SIMILARITY: BELONGS TO THE ISPH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



DR EMBL: AE004050; AAF85215.1; -  
 DR InterPro: IPR003451; LytB.  
 DR Pfam: PF02401; LytB; 1.  
 DR TIGRfams: TIGR00216; LytB; 1.  
 DR Isoprene biosynthesis; Complete proteome.  
 SQ SEQUENCE 316 AA; 34704 MW; 1A2E80B9A98D34A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 316;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDGVLA 223  
 |||||  
 Db 276 LVDGVLA 282

RESULT 42  
 ODPB\_BACSU STANDARD; PRT; 324 AA.

AC P21882;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Pyruvate dehydrogenase E1 component, beta subunit (EC 1.2.4.1) (S  
 GN PDHB OR ACEB.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=90368558; PubMed=1697575;  
 RA Hemila H., Palva A., Paulin L., Arvidsson S., Palva I.;  
 RT "Secretory S complex of Bacillus subtilis: sequence analysis and  
 RT identity to pyruvate dehydrogenase";  
 RL J. Bacteriol. 172:5052-5063(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97124187; PubMed=8969500;  
 RA Winters P., Caldwell R., Enfield L., Ferrari E.;  
 RT "The ampS-nprE (124 degrees-127 degrees) region of the Bacillus  
 RT subtilis 168 chromosome: sequencing of a 27 kb segment and  
 RT identification of several genes in the area";  
 RL Microbiology 142:3033-3037(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Caldwell R.M., Ferrari E.;  
 RT "Sequence analysis of the mobA-amps region of the Bacillus subtilis  
 RT chromosome";

Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall  
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains  
 CC multiple copies of three enzymatic components: pyruvate  
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and  
 CC lipoamide dehydrogenase (E3).  
 CC -1- FUNCTION: THE B. SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-CHAIN  
 CC 2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-  
 CC acetylhydrolipoamide + CO(2).  
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M57435; AAA62682.1; -  
 DR EMBL: AF012285; AAC24933.1; -  
 DR EMBL: Z59111; CAB13332.1; -  
 DR PIR: C36718; C36718.  
 DR HSSSP: P09061; I0S0.  
 DR Subtilist; BG10208; pdhb.  
 DR InterPro: IPR000360; Transketolase.  
 DR Pfam: PF02779; transket-pyr; 1.  
 DR Pfam: PF02780; transketolase\_C; 1.  
 DR Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;  
 KW Complete proteome.

FT INIT\_MET 0  
 SQ SEQUENCE 324 AA; 35343 MW; D2A7C9B32DED0F0D CRC64;

Query Match 2.5%; Score 7; DB 1; Length 324;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGIGL 65  
 |||||  
 Db 65 LALGIGL 71

RESULT 43  
 SRA6\_CAEEL

ID SRA6\_CAEEL STANDARD; PRT; 329 AA.

AC Q09208;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serpentine receptor class alpha 6 (Sra-6 protein).  
 GN SRA-6 OR AH6.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol NZ;  
 RA Tassal B.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE C. ELEGANS RECEPTOR-LIKE PROTEIN SRA  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z48009; CAA88083.1; -  
 DR WormPep; AH6.10; CE01451.  
 DR InterPro: IPR000344; Sra\_chemo\_Ce.  
 DR Pfam: PF02117; Sra; 1.  
 DR PRINTS: PR00697; TMPTROTEINRA.  
 KW Transmembrane; Multigene family.  
 FT TRANSMEM 26 46  
 FT TRANSMEM 104 124  
 FT TRANSMEM 143 163  
 FT TRANSMEM 187 207  
 FT TRANSMEM 238 258  
 FT TRANSMEM 273 293  
 SQ SEQUENCE 329 AA; 37951 MW; 7C0963ADA53A29F3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 329;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 134 RTRRAR 140  
 |||||

```

Db      297 RRTARR 303

RESULT 44
TAG2_RAT ID TAG2_RAT STANDARD; PRT; 335 AA.
AC 035777;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IMplantation-associated protein.
GN IAG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Uterus;
RA Barkai U., Amit A., Leising J.B., Kraicer P.F., Kidron T.;
RT "Implantation-associated uterine protein.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE OSF3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF008554; AAB63284.2;
DR InterPro: IPR000063; ThioRed.
KW Transmembrane.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 271 291 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
SQ SEQUENCE 335 AA; 37992 MW; AFB9DB35F1A06091 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 335;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LGLLAV 76
Db 192 LGLLAV 198

RESULT 45
LEU3_THEAO ID LEU3_THEAO STANDARD; PRT; 344 AA.
AC P24098;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-Isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEUB.
OS Thermus aquaticus.
OC Bacteria; Thermus/Delnococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_Taxid=271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YT1;
RX MEDLINE=92041736; Pubmed=1939005;
RA Kiriho H., Oshima T.;
RT "Molecular cloning and nucleotide sequence of 3-Isopropylmalate
RT dehydrogenase gene (leub) from an extreme thermophile, Thermus

```

```

RT aquaticus YT-1.";
RL J. Biochem. 109:852-857(1991).
CC -1- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
CC NAD(+) -> 3-carboxy-4-methyl-2-oxopentanoate + NADH.
CC -1- PATHWAY: Leucine biosynthesis; third step.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D10700; BAA01542.1;
DR PIR: JX0173; JX0173.
DR HSSP: P00351; 1XNA.
DR InterPro: IPR001804; Isodh.
DR InterPro: IPR004429; Leub.
DR Pfam: PF00180; Isodh; 1.
DR TIGRFAMs: TIGR00169; leub; 1.
DR PROSITE: PS00470; IDH_IMDH; 1.
KW Oxidoreductase; leucine biosynthesis; NAD.
SQ SEQUENCE 344 AA; 36949 MW; 4F31A9444E826408 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 SGLALR 254
Db 88 SGLALR 94

```

Search completed: May 8, 2003, 06:56:24  
Job time : 40 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 04:23:48 ; Search time 61 Seconds

(without alignments)  
620.380 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: A\_Geneseq\_101002:\*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	284	19	AAW47525 Homo sapiens tumou
2	249	87.7	249	20	AAW93369 Human tumour necro
3	249	87.7	249	21	AAW93369 Human tumour necro
4	249	87.7	249	23	AAW93338 Amino acid sequenc
5	249	87.7	249	23	AAW93338 Human PRO207 antit
6	241	84.9	249	19	AAW29745 TNF related endoth
7	241	84.9	249	22	AAW29745 Human TREPA (TNF r
8	207	72.9	273	22	AAW03499 Human TREPA (TNF r
9	146	51.4	146	22	AAW0895 Human TREPA (TNF r
10	143	50.4	189	19	AAW29746 TNF related endoth

11	143	50.4	189	22	AAW0892 Human UT41ag TREP
12	107	37.7	208	20	AAW93590 Human TREL3 protei
13	46	16.2	211	20	AAW93591 Mouse TREL3 protei
14	32	11.3	225	19	AAW47524 Mus musculus tumou
15	32	11.3	225	21	AAW07527 Amino acid sequenc
16	9	3.2	58	21	AAW01265 Human secreted pro
17	9	3.2	748	20	AAW01266 Human secreted pro
18	9	3.2	749	23	AAW73512 Extended amino aci
19	9	3.2	749	23	ABW73512 M vaccae GVS-9 pro
20	8	2.8	55	22	AAW51863 Propionibacterium
21	8	2.8	55	22	ABW42637 Peptide #10143 enc
22	8	2.8	55	22	ABW25983 Protein #7982 enco
23	8	2.8	55	22	AAW63528 Human brain expres
24	8	2.8	55	22	AAW76342 Human bone marrow
25	8	2.8	55	22	AAW20858 Peptide #7292 enco
26	8	2.8	55	22	AAW6451 Peptide #10488 enc
27	8	2.8	55	23	ABW45635 Human peptide enco
28	8	2.8	65	22	ABW43981 Peptide #11487 enc
29	8	2.8	65	22	ABW26890 Protein #8889 enco
30	8	2.8	65	22	AAW64984 Human brain expres
31	8	2.8	65	22	AAW7706 Human bone marrow
32	8	2.8	65	22	AAW21621 Peptide #8055 enco
33	8	2.8	65	22	AAW37923 Peptide #11960 enc
34	8	2.8	65	22	ABW46737 Novel human diagno
35	8	2.8	69	22	ABW03723 Novel human diagno
36	8	2.8	71	22	ABW03663 Novel human diagno
37	8	2.8	84	22	AAW61282 Propionibacterium
38	8	2.8	110	22	ABW20253 Novel human diagno
39	8	2.8	117	22	AAW08094 Human polypeptide
40	8	2.8	184	22	ABW67712 Drosophila melanog
41	8	2.8	190	23	ABP28041 Streptococcus poly
42	8	2.8	198	22	AAW23684 Human EST encoded
43	8	2.8	222	23	ABW41674 Human ovarian anti
44	8	2.8	266	22	ABW16279 Novel human diagno
45	8	2.8	307	22	ABW01186 Novel human diagno

## ALIGNMENTS

```
RESULT 1
AAW47525
ID AAW47525 standard; Protein: 284 AA.
AC AAW47525;
XX 21-JUL-1998 (first entry)
DT
XX Homo sapiens, tumour necrosis factor related ligand (TRELL).
DE
XX TRELL: tumour necrosis factor related ligand; tnfr treatment;
KW cancer; autoimmune disease; immune system; stimulation; suppression;
KW graft rejection.
XX
XX Homo sapiens.
XX OS
XX WO9805783-A1.
XX PD
XX 12-FEB-1998.
XX PF
XX 07-AUG-1997; 97WO-US13945.
XX PR
XX 18-MAR-1997; 97US-0040820.
XX PR
XX 07-AUG-1996; 96US-0023541.
XX PR
XX 18-OCT-1996; 96US-0028515.
XX PA
XX (BIOJ) BIOGEN INC.
XX (URGE-) UNIV GENOVA FACULTY MEDICINE.
XX Brownig JL, Chicheportliche Y;
XX WPI: 1998-145619/13.
XX N-PSDB: AAW18600.
```

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
XX  
PS Claim 12: Pages 50-51, 69pp: English.  
XX

Claim 12; Pages 50-51; 69pp; English.

The sequence is that of human tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon- $\gamma$ , to induce cell death or, treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. It's coding sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective amounts of a vector, e.g. a virus comprising a gene encoding TRELL. It may also be of use in the preparation of probe probes for screening natural/synthetic DNAs for TRELL-encoding sequences and for antisense therapy.

**SQ Sequence 284 AA;**

Query Match	100.0%	Score 284;	DB 19;	Length 284;
Best Local Similarity	100.0%	Pred. No. 2,26-252;		
Matches 284;	Conservative 0;	Mismatches 0;	Gaps 0;	

QY	1	MSLDFEELSARRLPLPSSLSGSDGCAVRAQOPAPMAPMAARSRQRRRGRGEGTALLVPLA	60
Db	1	MSLDFEELSARRLPLPSSLSGSDGCAVRAQOPAPMAPMAARSRQRRRGRGEGTALLVPLA	60
QY	61	LGGLGLACGLLLAVVSLGSRASLSAOEPAOEELVAEEDDPSBELNPQTEESODPAPFL	120
Db	61	LGGLGLACGLLLAVVSLGSRASLSAOEPAOEELVAEEDDPSBELNPQTEESODPAPFL	120
QY	121	NRLVRPERSAPKGGKTRABARILAAHYEVHPRGDDGAQVDTGVSQMEARINSSPLR	180
Db	121	NRLVRPERSAPKGGKTRABARILAAHYEVHPRGDDGAQVDTGVSQMEARINSSPLR	180
QY	181	YNRQIGFEIVTRAGLYLYLYCOVNHDECKAAVYLKLDLLVYDVIALRCLEEFESATPAASSLGP	240
Db	181	YNRQIGFEIVTRAGLYLYLYCOVNHDECKAAVYLKLDLLVYDVIALRCLEEFESATPAASSLGP	240
QY	241	QLRLCOVSGLLARPGSSLLRPLPMANHLAAAPLTYFGGLFYQNH	284
Db	241	QLRLCOVSGLLARPGSSLLRPLPMANHLAAAPLTYFGGLFYQNH	284

## RESULT 2

ID	standard; Protein; 249 AA
AA09369	

AC AAY09369;

DT 15-JUL-1999 (first entry)

Human tumour necrosis factor Apo-3 ligand protein sequence

KM Human; tumour necrosis factor; Apo-3 ligand; Lymphotoxin; apoptosis  
KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
KM cancer.

OS	Homo sapiens.
XX	
PN	W09919490-A1.
XX	
PD	22-APR-1999.
XX	

PF	09-OCT-1998;	98WO-US21407
XX		
PR	17-DEC-1997;	97US-0069862
PR	10-OCT-1997;	97US-0062037
XX		

PA (GETH ) GENENTECH INC  
YY

PL Asnkenazi AU, Marsters SA, Pittl R,  
XY

DA WFL; 1999-2019/24  
DR N-PSDB: AAX56000

21.16  
PT  
New human Apo-3

PS Claim 1: Fig 1: 74pb: English.

CC The present sequence represents

CC The present sequence represents a human tumour necrosis factor (TNF)  
CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
CC cytosolic activity. Apo-3 ligand can be used to induce apoptosis in  
CC mammalian cancer cells, to induce NF-kappa-dependent transcription and  
CC to induce JNK/SAPK-dependent responses in mammalian cells.

**SQ Sequence 249 AA;**

Query Match	87.7%	Score 249	DB 20	Length 249
Best Local Similarity	100.0%	Pred. NO.	2,76-220	
Matches 249	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	56	MAAKRSVQRKRKRREBPGLLVLPALGKUALGUALGGLLVAVSGSRASISADBPABEEL	95
Db	1	MAARSRORRKRREBPGLLVLPALGKUALGUALGGLLVAVSGSRASISADBPABEEL	60
QY	96	VAEEDDOPSELNPOTIESODPAPRLNLRVPRRSAPGRKTRARRALAAHYEVNPRGOD	155
Db	61	VAEEDDOPSELNPOTIESODPAPRLNLRVPRRSAPGRKTRARRALAAHYEVNPRGOD	120
QY	156	GAGAGVGTGVSQWEARINSSSPRYNRQIGEFVYTSAGLYYLVCQVNHDEGAAVYTKLD	215
Db	121	GAGAGVGTGVSQWEARINSSSPRYNRQIGEFVYTSAGLYYLVCQVNHDEGAAVYTKLD	180
QY	216	LVNDSYALACIEEFNSTAASSSLGAPQLRLCOVSGSLLRPGSSLRITRLPMHLKAAPFL	275
Db	181	LVNDSYALACIEEFNSTAASSSLGAPQLRLCOVSGSLLRPGSSLRITRLPMHLKAAPFL	240
QY	276	TYRGLFQVH 284 .	
Db	241	TYRGLFQVH 249	

## RESULT 3

ID AAB07526 standard; protein; 249 AA

AC AAB07526

DT 20-OCT-2000 (first entry)

Amino acid sequence of a soluble recombinant human TWEAK protein

KM TWEAK protein; immunological disorder; immune response; inflammation;  
KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
KM Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour

OS	Homo sapiens.
XX	
PN	WO2000042073-A1

PD	20-JUL-2000.
XX	
PF	14-JAN-2000; 2000WO-US01044

PR 15-JAN-1999; 99US-0116168

PA (BIOJ ) BIOGEN INC.  
 XX  
 XX  
 PI Remmert P;  
 XX  
 DR WPI; 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 XX Disclosure; Fig 1; 45pp; English.  
 XX  
 XX The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, graft-versus-host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune  
 CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 CC  
 XX Sequence 249 AA;  
 SQ  
 Query Match 87.7%; Score 249; DB 21; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 2,7e-220; Mismatches 0; Gaps 0;  
 Matches 249; Conservative 0; Indels 0;  
 OY 36 MAARSQRRRGRGEPGATLVPALGLALACGLLAVYSLGRASLSAOEPQDEL 95  
 DB 1 MAARSQRRRGRGEPGATLVPALGLALACGLLAVYSLGRASLSAOEPQDEL 60  
 OY 96 VAEDDQDSELPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEYHPRPGD 155  
 DB 61 VAEDDQDSELPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEYHPRPGD 120  
 OY 156 GAOAGVDGVSQWEERINSSSLRNQIGEFITRRLVLYLCQVHDESKAVYLKLD 215  
 DB 121 GAOAGVDGVSQWEERINSSSLRNQIGEFITRRLVLYLCQVHDESKAVYLKLD 180  
 OY 216 LLYDGVLAIRCLSEFSATNAASLGLPOLRLCOVSGLLALRPGSSLRITLPMALTKAPPL 275  
 DB 181 LLYDGVLAIRCLSEFSATNAASLGLPOLRLCOVSGLLALRPGSSLRITLPMALTKAPPL 240  
 OY 276 TYFGLFQVH 284  
 DB 241 TYFGLFQVH 249  
 RESULT 4  
 ID AAY95338  
 AAY95338 standard; Protein; 249 AA.  
 AC AAY95338;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 antitumour protein.  
 XX  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide 1..40  
 FT /label= Signal\_peptide  
 FT Protein 41..249  
 FT /label= PRO207

FT Modified-site 27..33  
 FT /note= "N-myristoylation"  
 FT 29..35  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 36..42  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 45..51  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 118..124  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 121..127  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 125..131  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 128..134  
 FT Modified-site  
 FT /note= "N-myristoylation"  
 FT 139..143  
 FT Modified-site  
 FT /note= "Asn is N-glycosylated"  
 FT 10..14  
 FT Modified-site  
 FT /note= "amidation"  
 FT 97..101  
 FT Modified-site  
 FT /note= "amidation"  
 FT 24..35  
 FT Peptide  
 FT /note= "prokaryotic membrane lipid protein lipid"  
 XX  
 PN W0200037638-A2.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99WO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130322.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pittl RM, Wood WT;  
 PI  
 XX WPI; 2000-442668/38.  
 DR N-PSDB; AAA49717.  
 XX  
 DR Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO309 or  
 PT PRO866  
 XX  
 PS Claim 19; Fig 4; 172pp; English.  
 XX  
 CC The present sequence is that of human antitumour protein PRO207,  
 CC as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207  
 CC shows amino acid sequence identity to tumour necrosis factor family  
 CC members, especially human lymphotoxin-beta (23.4%) and human CD40  
 CC ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting  
 CC the growth of a tumour cell comprises exposing the tumour cell  
 CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,  
 CC PRO326, PRO362, PRO509 or PRO866 (see AAY95337-49), their  
 CC agonists or chimeric polypeptides incorporating them. The tumour  
 CC is especially a cancer selected from breast, ovarian, renal,  
 CC colorectal, uterine, prostate, lung, bladder and central nervous  
 CC system cancer, melanoma and leukaemia. Methods for the recombinant  
 CC expression of the antitumour proteins are also provided.  
 CC  
 XX Sequence 249 AA;  
 SQ  
 Query Match 87.7%; Score 249; DB 21; Length 249;



DR WPI: 1998-447255/38.  
 DR N-PSDB: AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC R1b0zymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 SQ Sequence 249 AA:  
 Query Match 84.9%; Score 241; DB 19; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213; Indels 0; Gaps 0;  
 Matches 241; Conservative 0; Mismatches 0;  
 QY 44 RRGRRGEPFTALLVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 103  
 DB 9 RRGRRGEPFTALLVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 68  
 QY 104 SELNPQTEESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOGAAGVND 163  
 DB 69 SELNPQTEESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOGAAGVND 128  
 QY 164 TVSGWEARINSSPLRYNRQIGEFVTYRAGLYLYCQVHPEDEKRAVYIKDLLVDGYLA 223  
 DB 129 TVSGWEARINSSPLRYNRQIGEFVTYRAGLYLYCQVHPEDEKRAVYIKDLLVDGYLA 188  
 QY 224 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGELFV 283  
 DB 189 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGELFV 248  
 QY 284 H 284  
 DB 249 H 249  
 RESULT 7  
 AAE00891  
 ID AAE00891 standard; Protein: 249 AA.  
 XX  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulnerability.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 FT /label= Extracellular\_domain  
 XX  
 PN US6207642-B1.

PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB: AAD04350.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 XX  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 SQ Sequence 249 AA:  
 Query Match 84.9%; Score 241; DB 22; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213; Indels 0; Gaps 0;  
 Matches 241; Conservative 0; Mismatches 0;  
 QY 44 RRGRRGEPFTALLVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 103  
 DB 9 RRGRRGEPFTALLVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 68  
 QY 104 SELNPQTEESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOGAAGVND 163  
 DB 69 SELNPQTEESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOGAAGVND 128  
 QY 164 TVSGWEARINSSPLRYNRQIGEFVTYRAGLYLYCQVHPEDEKRAVYIKDLLVDGYLA 223  
 DB 129 TVSGWEARINSSPLRYNRQIGEFVTYRAGLYLYCQVHPEDEKRAVYIKDLLVDGYLA 188  
 QY 224 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGELFV 283  
 DB 189 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGELFV 248  
 QY 284 H 284  
 DB 249 H 249  
 RESULT 8  
 AAU03499  
 ID AAU03499 standard; Protein: 273 AA.  
 XX  
 AC AAU03499;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;

KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KW fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN MO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 19-DEC-2000; 2000MO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-417975/44.  
 DR N-PSDB: AAS03964.  
 XX  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis; e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 XX  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-L2-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC hemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 CC  
 SQ Sequence 273 AA;  
 XX  
 XX  
 Query Match 72.9%; Score 207; DB 22; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-181;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 78 SLGRASLSLQAEPAQELVAEEDDDPSLNPOTEESSODPAFLRLVPRRSAPKGRKTR 137  
 DB 67 SLGRASLSLQAEPAQELVAEEDDDPSLNPOTEESSODPAFLRLVPRRSAPKGRKTR 126  
 QY 138 ARRAIAAHYEVHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGLYYL 197  
 DB 127 ARRAIAAHYEVHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGLYYL 186  
 QY 198 LYCOVHDEGKAVYLLKIDLLVDGVALRCLDEFSATASSLGPOLRLCOVSGLLALRPGS 257  
 DB 187 LYCOVHDEGKAVYLLKIDLLVDGVALRCLDEFSATASSLGPOLRLCOVSGLLALRPGS 246  
 QY 258 SLRIRTLFWAHLKAAAPFLTYTGFQVH 284

DB 247 SLRIRTLFWAHLKAAAPFLTYTGFQVH 273  
 RESULT 9  
 AAE00895  
 ID AAE00895 standard; Protein: 146 AA.  
 XX  
 AC AAE00895;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent) fragment.  
 XX  
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vunerary.  
 XX  
 OS Homo sapiens.  
 OS  
 PN US6207642-B1.  
 XX  
 PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX  
 PS Example 14; Fig 1; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is human TREPA fragment.  
 CC  
 SQ Sequence 146 AA;  
 XX  
 XX  
 Query Match 51.4%; Score 146; DB 22; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-126;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 139 RRAIAAHYEVHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGLYYL 198  
 DB 1 RRAIAAHYEVHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGLYYL 60  
 QY 199 YCOVHDEGKAVYLLKIDLLVDGVALRCLDEFSATASSLGPOLRLCOVSGLLALRPGS 258  
 DB 61 YCOVHDEGKAVYLLKIDLLVDGVALRCLDEFSATASSLGPOLRLCOVSGLLALRPGS 120  
 QY 259 LRIRTLFWAHLKAAAPFLTYTGFQVH 284  
 DB 121 LRIRTLFWAHLKAAAPFLTYTGFQVH 146  
 RESULT 10  
 AAM29746  
 ID AAM29746 standard; Protein: 189 AA.



XX AAM29746;  
 AC  
 XX 27-OCT-1998 (first entry)  
 DT  
 XX  
 DE TNF related endothelium proliferative agent protein 2.  
 DE  
 XX TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 KW  
 OS Homo sapiens.  
 OS  
 PN W09835061-A2.  
 PN  
 XX 13-AUG-1998.  
 PD  
 XX 12-FEB-1998; 98WO-US02859.  
 PF  
 XX 10-FEB-1998; 98US-0021706.  
 PR  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 PA  
 XX WILEY SR;  
 PI  
 XX WPI; 1998-447255/38.  
 DR  
 XX Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 PT  
 XX Claim 16; Page 125-6; 142pp; English.  
 PS  
 XX The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX Sequence 189 AA;  
 SQ  
 Query Match 50.4%; Score 143; DB 19; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-123;  
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 IAAHEVHPRGQDGAQGVDTGVSQWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 201  
 |||||||  
 DB 47 IAAHEVHPRGQDGAQGVDTGVSQWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 106  
 |||||||  
 QY 202 VHFDEGKAVYLLKLDLVGVTLALRCLSEFSATPAASLGPOLRLCOVSGLLALRPSSIRI 261  
 |||||||  
 DB 107 VHFDEGKAVYLLKLDLVGVTLALRCLSEFSATPAASLGPOLRLCOVSGLLALRPSSIRI 166  
 |||||||  
 QY 262 RTLPMAHKAAAPFLTYFGLFOVH 284  
 |||||||  
 DB 167 RTLPMAHKAAAPFLTYFGLFOVH 189  
 |||||||

RESULT 11  
 AAE00892  
 ID AAE00892 standard; Protein: 189 AA.  
 AC  
 XX AAE00892;  
 XX

DT 04-JUL-2001 (first entry)  
 XX  
 DE Human ULAflag TREPA soluble construct.  
 DE  
 XX Human; tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;  
 KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;  
 KW vulneryary; HUVEC; human umbilical vein endothelial cell; ULAflag.  
 KW  
 OS Homo sapiens.  
 OS  
 PN US6207642-B1.  
 PN  
 XX 27-MAR-2001.  
 PD  
 XX 26-JUN-1998; 98US-0105343.  
 PF  
 XX 12-FEB-1997; 97US-0798692.  
 PR  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 PA  
 XX WILEY SR;  
 PI  
 XX WPI; 2001-280760/29.  
 DR  
 XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX Example 2; Column 75-78; 53pp; English.  
 PS  
 XX The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts. The present  
 CC amino acid sequence is human ULAflag TREPA soluble construct. This  
 CC sequence which is a biologically active molecule is capable of inducing  
 CC proliferation in HUVEC (human umbilical vein endothelial cells).  
 CC  
 XX Sequence 189 AA;  
 SQ  
 Query Match 50.4%; Score 143; DB 22; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-123;  
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 IAAHEVHPRGQDGAQGVDTGVSQWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 201  
 |||||||  
 DB 47 IAAHEVHPRGQDGAQGVDTGVSQWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 106  
 |||||||  
 QY 202 VHFDEGKAVYLLKLDLVGVTLALRCLSEFSATPAASLGPOLRLCOVSGLLALRPSSIRI 261  
 |||||||  
 DB 107 VHFDEGKAVYLLKLDLVGVTLALRCLSEFSATPAASLGPOLRLCOVSGLLALRPSSIRI 166  
 |||||||  
 QY 262 RTLPMAHKAAAPFLTYFGLFOVH 284  
 |||||||  
 DB 167 RTLPMAHKAAAPFLTYFGLFOVH 189  
 |||||||

RESULT 12  
 AAM93590  
 ID AAM93590 standard; Protein: 208 AA.  
 AC  
 XX AAM93590;  
 AC  
 XX 18-JUN-1999 (first entry)  
 DT  
 XX Human TNRL3 protein.  
 DE  
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KW

KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW Ap06; Ap08; Ap09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 PN W09911791-A2.  
 PD  
 PD 11-MAR-1999.  
 XX  
 XX  
 PF 04-SEP-1998; 98WC-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 PI Chaudhary PM;  
 XX  
 XX WPI; 1999-205191/17.  
 DR N-PSDB; AAX23424.  
 DR  
 XX  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PS  
 PS Claim 40; Fig 13A; 156pp; English.  
 XX  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: Ap04, Ap06, Ap08 and Ap09 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. Ap04 is useful for diagnosing prostate cancer  
 CC by determining levels of Ap04 in an individual. Prostate cancer can also  
 CC be treated using Ap04 selective binding agents linked to a therapeutic  
 CC moiety. Ap04 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. Ap04 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in Ap04  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using Ap04 polypeptides/active  
 CC fragments and Ap04 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of Ap04 and detecting a change in level of Ap04  
 CC activity. The method is performed in vivo or in vitro. Ap0 polypeptides  
 CC are all useful as immunogens for preparing antibodies. Ap04 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. Ap08 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 XX  
 SQ Sequence 208 AA;  
 Query Match 37.7%; Score 107; DB 20; Length 208;  
 Best Local Similarity 99.5%; Pred. No. 5.8e-90;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 77 VSLGSRSLAQEPAGEELVAEEDQDPSELNPTQTESQDPAFNLVRRSPKGRKT 136  
 DB 1 VSLGSRSLAQEPAGEELVAEEDQDPSELNPTQTESQDPAFNLVRRSPKGRKT 60  
 QY 137 RARRAIAHYEVHPRGDDAAGVDDTSGWGEARINSSPLLYNQIGFTYTRAGLY 196  
 DB 61 RARRAIAHYEVHPRGDDAAGVDDTSGWGEARINSSPLLYNQIGFTYTRAGLY 120  
 QY 197 YLYCOVHFDEGKAVYLRDLVDVLAIRCLIEFSATASLIGPOLRCQVSGLLALRP 256  
 DB 121 YLYCOVHFDEGKAVYLRDLVDVLAIRCLIEFSATASLIGPOLRCQVSGLLALRP 180  
 QY 257 SSLRIRLTPWAHLKAPFLTYTGLQVH 284  
 DB 181 SSLRIRLTPWAHLKAPFLTYTGLQVH 208

RESULT 13  
 AAW93591  
 ID AAW93591 standard; Protein: 211 AA.  
 XX  
 XX  
 AC AAW93591;  
 XX  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 XX  
 DE Mouse TNRL3 protein.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; Ap04;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW Ap06; Ap08; Ap09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse.  
 XX  
 XX  
 OS Mus sp.  
 XX  
 XX  
 PN W09911791-A2.  
 PD  
 PD 11-MAR-1999.  
 XX  
 XX  
 PF 04-SEP-1998; 98WC-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 PI Chaudhary PM;  
 XX  
 XX WPI; 1999-205191/17.  
 DR N-PSDB; AAX23425.  
 DR  
 XX  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PS  
 PS Claim 40; Fig 13B; 156pp; English.  
 XX  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: Ap04, Ap06, Ap08 and Ap09 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. Ap04 is useful for diagnosing prostate cancer  
 CC by determining levels of Ap04 in an individual. Prostate cancer can also  
 CC be treated using Ap04 selective binding agents linked to a therapeutic  
 CC moiety. Ap04 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. Ap04 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in Ap04  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using Ap04 polypeptides/active  
 CC fragments and Ap04 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of Ap04 and detecting a change in level of Ap04  
 CC activity. The method is performed in vivo or in vitro. Ap0 polypeptides  
 CC are all useful as immunogens for preparing antibodies. Ap04 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. Ap08 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 XX  
 SQ Sequence 211 AA;  
 Query Match 16.2%; Score 46; DB 20; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-34;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 239 GPOLRCQVSGLLALRPSSSLRIRLTPWAHLKAPFLTYTGLQVH 284  
 DB 166 GPOLRCQVSGLLALRPSSSLRIRLTPWAHLKAPFLTYTGLQVH 211

AAW47524  
ID AAW47524 standard; Protein; 225 AA.  
XX  
AC AAW47524;  
XX  
DT 21-JUL-1998 (first entry)  
XX  
DE Mus musculus tumour necrosis factor related ligand (TRELL).  
XX  
KW TRELL; tumour necrosis factor related ligand; tnf; treatment;  
KM cancer; autoimmune disease; immune system; stimulation; suppression;  
XX graft rejection.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..21  
FT /note= "hydrophobic, transmembrane domain"  
XX  
PN WO9805783-A1.  
XX  
PD 12-FEB-1998.  
XX  
PF 07-AUG-1997; 97WO-US13945.  
XX  
PR 18-MAR-1997; 97US-0040820.  
PR 07-AUG-1996; 96US-0023541.  
PR 18-OCT-1996; 96US-0028515.  
XX  
PA (BIOJ ) BIOGEN INC.  
PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
XX  
PI Browning JL, Chicheportiche Y;  
XX  
DR WPI; 1998-145619/13.  
DR N-PSDB; AAV18599.  
XX  
PT Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
PS Claim 12; Pages 48-50; 69pp; English.  
XX  
XX The sequence is that of mouse tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. It's coding sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of prepare probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.  
XX  
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 170  
|||||  
DB 80 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 111

RESULT 15  
AAB07527  
ID AAB07527 standard; protein; 225 AA.  
XX  
AC AAB07527;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
XX  
KW TWEAK protein; immunological disorder; immune response; inflammation;  
KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
XX Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
XX  
OS Mus sp.  
XX  
PN WO200042073-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 14-JAN-2000; 2000WO-US01044.  
XX  
PR 15-JAN-1999; 99US-0116168.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX  
PI Remmert P;  
XX  
DR WPI; 2000-476036/41.  
XX  
PT Preventing and treating immune responses using modulators, especially  
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
PT treating e.g. inflammation and graft versus host disease -  
XX  
XX Disclosure: Fig 1; 45pp; English.  
XX  
PS The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.  
XX  
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 21; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 170  
|||||  
DB 80 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 111

RESULT 16  
AAG01265  
ID AAG01265 standard; Protein; 58 AA.  
XX  
AC AAG01265;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein, SEQ ID NO: 5346.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping.  
XX  
OS Homo sapiens.

XX EP1033401-A2.  
 PN 06-SEP-2000.  
 PD 21-FEB-2000; 2000EP-0200610.  
 PE 26-FEB-1999; 99US-0122487.  
 PR (GENSET).  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI: 2000-500381/45.  
 DR N-PSDB: AAC01271.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 5346; 71pp + CD-ROM; English.  
 CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 CC  
 SQ Sequence 58 AA;  
 QY Query Match 3.2%; Score 9; DB 21; Length 58;  
 DB Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MSLDFEIS 9  
 DB 46 MSLDFEIS 54  
 AC AAG01266;  
 DE 06-OCT-2000 (first entry)  
 DE Human secreted protein, SEQ ID NO: 5347.  
 DE Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX Homo sapiens.  
 OS  
 XX EPI033401-A2.  
 PN 06-SEP-2000.  
 PD 21-FEB-2000; 2000EP-0200610.  
 PE 26-FEB-1999; 99US-0122487.  
 PR (GENSET).  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI

XX WPI: 2000-500381/45.  
 DR N-PSDB: AAC01272.  
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 5347; 71pp + CD-ROM; English.  
 CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 CC  
 SQ Sequence 58 AA;  
 QY Query Match 3.2%; Score 9; DB 21; Length 58;  
 DB Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MSLDFEIS 9  
 DB 46 MSLDFEIS 54  
 AC AAY14906;  
 DE 25-OCT-1999 (first entry)  
 DE Extended amino acid sequence for GVs-9.  
 DE Mycobacterium vaccae protein; antigen; T cell activation; cytokine;  
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;  
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;  
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;  
 KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;  
 KW squamous cell carcinoma; melanoma.  
 XX Mycobacterium vaccae.  
 OS  
 XX WO9932634-A2.  
 PN 01-JUL-1999.  
 PD 23-DEC-1998; 98WO-NZ00189.  
 PE 04-DEC-1998; 98US-0205426.  
 PR 23-DEC-1997; 97US-0996624.  
 PR 23-DEC-1997; 97US-0997080.  
 PR 23-DEC-1997; 97US-0997362.  
 PR 11-JUN-1998; 98US-0095855.  
 PR 17-SEP-1998; 98US-0156181.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;  
 PI WPI: 1999-430163/36.  
 DR N-PSDB: AAZ11368.

XX Enhancing immune response to an antigen  
 PT  
 XX  
 PS Claim 1; Page 209-210; 243pp; English.  
 XX  
 CC The invention provides heat-killed *Mycobacterium vaccae*, or recombinant  
 CC *M. vaccae* proteins. The *M. vaccae* proteins may be employed to activate  
 CC T cells and natural killer cells, to stimulate the production of  
 CC cytokines, to enhance the expression of co-stimulatory molecules on  
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation  
 CC and function. The proteins can be expressed by standard recombinant  
 CC methodology. Pharmaceutical compositions comprising the proteins or  
 CC nucleic acid sequences encoding the proteins can be used for the  
 CC treatment, prevention, and detection of disorders including infections  
 CC diseases, immune disorders and cancer. In particular, the compounds and  
 CC methods are used for treatment of diseases of the respiratory system,  
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,  
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as  
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,  
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell  
 CC carcinoma and melanoma.  
 SQ Sequence 748 AA:  
 Query Match 3.2%; Score 9; DB 20; Length 748;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGLGLALA 68  
 |||||  
 DB 282 ALGLGLALA 290  
 RESULT 19  
 ABB73512  
 ID ABB73512 standard; Protein; 749 AA.  
 AC ABB73512;  
 XX  
 XX 08-APR-2002 (first entry)  
 DT  
 XX  
 DE M vaccae GVS-9 protein SEQ ID NO: 154.  
 XX  
 XX Skin disorder: psoriasis; atopic dermatitis; allergic contact dermatitis;  
 KW alopecia areata; skin cancer; *Mycobacterium vaccae*; melanoma; cytostatic;  
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;  
 KW Th2 immune response; immunomodulatory.  
 XX  
 OS *Mycobacterium vaccae*.  
 XX  
 XX US6328978-B1.  
 PN  
 XX 11-DEC-2001.  
 PD  
 XX 02-JUN-1999; 99US-0324542.  
 PF  
 XX 23-DEC-1997; 97US-0997080.  
 PR  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA  
 PI Watson JD, Tan PLJ, Prestidge R;  
 XX  
 XX WPI: 2002-138361/18.  
 DR N-PSDB; ABI36274.  
 XX  
 XX Inhibiting skin inflammation associated with skin disorder e.g.  
 PT psoriasis, by administering composition comprising delipidated and  
 PT deglycolipidated *Mycobacterium vaccae* cells or *Mycobacterium vaccae*  
 PT culture filtrate  
 XX  
 XX Example 6; Column 161-164; 116pp; English.  
 PS  
 PS The present invention relates to a method of inhibiting skin inflammation  
 CC

CC associated with a skin disorder selected from psoriasis, atopic  
 CC dermatitis and allergic contact dermatitis, which involves administering  
 CC a composition containing delipidated and deglycolipidated *Mycobacterium*  
 CC *vaccae* cells or *M. vaccae* culture filtrate. The skin disorder to be  
 CC treated may also include alopecia areata, and skin cancers such as basal  
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition  
 CC acts by inhibiting the Th2 immune response. The present sequence is a  
 CC protein described in the exemplification of the invention.  
 SQ Sequence 749 AA:  
 Query Match 3.2%; Score 9; DB 23; Length 749;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGLGLALA 68  
 |||||  
 DB 282 ALGLGLALA 290  
 RESULT 20  
 AAU51863  
 ID AAU51863 standard; Protein; 54 AA.  
 AC AAU51863;  
 XX  
 XX 27-FEB-2002 (first entry)  
 DT  
 XX  
 DE Propionibacterium acnes immunogenic protein #12759.  
 XX  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 XX WO200181581-A2.  
 PN  
 XX 01-NOV-2001.  
 PD  
 XX 20-APR-2001; 2001WO-US12865.  
 PF  
 XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 XX WPI: 2001-616774/71.  
 DR N-PSDB; AAS59552.  
 XX  
 XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 PT  
 PS Example 1; SEQ ID NO 13058; 1069pp; English.  
 PS  
 XX Sequences AAU39105-AAU66017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 54 AA:

Query Match 2.88; Score 8; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 RRLPLPRS 18  
 |||||||  
 DB 21 RRLPLPRS 28

## RESULT 21

ABBA2637  
 ID ABB42637 standard; Peptide: 55 AA.

XX ABB42637;

DT 04-FEB-2002 (first entry)

DE Peptide #10143 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver.

PS Claim 27; SEQ ID NO 35272; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 55 AA:

Query Match 2.88; Score 8; DB 22; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68  
 |||||||  
 DB 12 LGIGIALA 19

## RESULT 22

ABB25983  
 ID ABB25983 standard; Protein: 55 AA.

XX ABB25983;

DT 23-JAN-2002 (first entry)

DE Protein #7982 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT hearts.

PS Claim 15; SEQ ID NO 27753; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 55 AA:

Query Match 2.88; Score 8; DB 22; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68  
 |||||||  
 DB 12 LGIGIALA 19

## RESULT 23

```

AAM63528
ID AAM63528 standard; Protein: 55 AA.
XX
AC AAM63528;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35633.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 35633; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 55 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGLGLALA 68
DB 12 LGLGLALA 19
XX
RESULT 24
AAM76342
ID AAM76342 standard; Protein: 55 AA.
XX
AC AAM76342;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36648.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.

```

```

XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 36648; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 55 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGLGLALA 68
DB 12 LGLGLALA 19
XX
RESULT 25
AAM20858
ID AAM20858 standard; Protein: 55 AA.
XX
AC AAM20858;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #17292 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

```

PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488901/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
XX  
PS Claim 27; SEQ ID No 25684; 487pp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENP; see A110068-A128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
XX  
SQ Sequence 55 AA;  
Query Match 2.8%; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 LGGLALAA 68  
| | | | | | | | | |  
DB 12 LGGLALAA 19  
RESULT 26  
AAM36451  
ID AAM36451 standard; Protein; 55 AA.  
XX  
XX AAM36451;  
AC  
XX  
DT 17-OCT-2001 (first entry)  
XX  
XX Peptide #10488 encoded by probe for measuring placental gene expression.  
DE  
XX  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157272-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US00663.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
PR  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX  
XX WPI; 2001-48897/53.  
DR  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX  
PS Claim 27; SEQ ID No 36720; 654pp; English.

XX  
XX The present invention relates to single exon nucleic acid probes (SENP;  
CC see A113315-A1157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
XX  
SQ Sequence 55 AA;  
Query Match 2.8%; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 LGGLALAA 68  
| | | | | | | | | |  
DB 12 LGGLALAA 19  
RESULT 27  
ABG45635  
ID ABG45635 standard; Peptide; 55 AA.  
XX  
XX  
XX ABG45635;  
AC  
XX  
XX 19-AUG-2002 (first entry)  
DT  
XX  
XX Human peptide encoded by genome-derived single exon probe SEQ ID 35300.  
DE  
XX  
XX Human; single exon probe; asthma; lung cancer; COPD; IID;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoids;  
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
XX primary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
OS  
XX  
XX Homo sapiens.  
XX  
XX  
XX WO200186003-A2.  
PN  
XX  
XX 15-NOV-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US00665.  
PF  
XX  
XX 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
PR  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX  
XX WPI; 2002-114183/15.  
DR  
XX  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
XX  
PS Claim 27; SEQ ID No 35300; 634pp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridize at high stringency to a



CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray, assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 55 AA:

Query Match 2.8%; Score 8; DB 23; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
| | | | | | | |  
DB 12 LGLGLALA 19

RESULT 28

ID ABB43981 standard; Peptide: 65 AA.

AC ABB43981;

DT 04-FEB-2002 (first entry)

DE Peptide #11487 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -

PS Claim 27; SEQ ID NO 36616; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 65 AA:

Query Match 2.8%; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
| | | | | | | |  
DB 22 LGLGLALA 29

RESULT 29

ID ABB26890 standard; Protein: 65 AA.

AC ABB26890;

DT 23-JAN-2002 (first entry)

DE Protein #8889 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

PS Claim 15; SEQ ID NO 28660; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 65 AA;

Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
XXXXXXXXXXXXXXXXXXXX  
DB 22 LGIGLALA 29

RESULT 30

ID AAM64984 standard; Protein: 65 AA.

XX  
AC AAM64984;

XX  
DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37089.

XX  
XX Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX  
XX epilepsy; cancer.

OS Homo sapiens.

XX  
PN MO200157275-A2.

XX  
PD 09-AUG-2001.

XX  
PF 30-JAN-2001; 2001WO-US00667.

XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
PR 04-OCT-2000; 2000GB-0024263.

XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
PI WPI; 2001-483446/52.

XX  
XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
XX brains -  
PS Example 4; SEQ ID NO: 37089; 650bp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention.

SQ Sequence 65 AA;

Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
XXXXXXXXXXXXXXXXXXXX  
DB 22 LGIGLALA 29

RESULT 31

ID AAM77706 standard; Protein: 65 AA.

XX  
AC AAM77706;

XX  
DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38012.

XX  
XX Human: bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma.

XX  
XX Homo sapiens.

XX  
PN MO200157276-A2.

XX  
PD 09-AUG-2001.

XX  
PF 30-JAN-2001; 2001WO-US00668.

XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
PR 04-OCT-2000; 2000GB-0024263.

XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
PI WPI; 2001-488900/53.

XX  
XX

PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human bone marrow -

XX  
PS Example 4; SEQ ID NO: 38012; 658bp + Sequence Listing; English.

XX  
CC The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is a  
XX protein encoded by one of the probes of the invention.

XX  
SQ Sequence 65 AA;

Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
XXXXXXXXXXXXXXXXXXXX  
DB 22 LGIGLALA 29

RESULT 32

ID AAM21621 standard; Protein: 65 AA.

XX  
XX

AC AAM21621;  
 XX  
 DN 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #8055 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID No 26447; 487bp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 65 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 61 LGIGLALA 68  
 |||||||  
 DB 22 LGIGLALA 29  
 RESULT 33  
 AAM37923  
 ID AAM37923 standard; Protein; 65 AA.  
 XX  
 AC AAM37923;  
 XX  
 DN 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #11960 encoded by probe for measuring placental gene expression.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX

XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27; SEQ ID No 38192; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see AAI33315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 CC  
 XX  
 SQ Sequence 65 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 61 LGIGLALA 68  
 |||||||  
 DB 22 LGIGLALA 29  
 RESULT 34  
 ABG46737  
 ID ABG46737 standard; Peptide; 65 AA.  
 XX  
 AC ABG46737;  
 XX  
 DN 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 36402.  
 XX  
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberculous sclerostis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00665.  
 XX  
 PR 04-FEB-2000; 2000US-180312P.  
 XX

PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000US-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2002-114183/15.  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 PS Claim 27; SEQ ID No 36402; 634pp; English.  
 XX  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridize at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridization of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridization to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Rudlak syndrome, sarcoidosis, pulmonary  
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WPI at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SO Sequence 65 AA:  
 Query Match 2.8%; Score 8; DB 23; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 LGIGLALA 68  
 DB 22 LGIGLALA 29  
 RESULT 35  
 ID ABG03723 standard; Protein; 69 AA.

XX  
 AC ABG03723;  
 XX 13-FEB-2002 (first entry)  
 DT Novel human diagnostic protein #3714.  
 DE  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KM food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS  
 PN WO200175067-A2.  
 PN  
 PD 11-OCT-2001.  
 PD  
 XX 30-MAR-2001; 2001MO-US08631.  
 PF  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 PR (HYSE-) HYSQ INC.  
 PA  
 PI Dmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS67910.  
 DR  
 XX  
 PS Claim 20; SEQ ID No 34082; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WPI  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SO Sequence 69 AA:  
 Query Match 2.8%; Score 8; DB 22; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 LGIGLALA 68  
 DB 16 LGIGLALA 23  
 RESULT 36  
 ID ABG03663 standard; Protein; 71 AA.  
 AC ABG03663;

XX 13-FEB-2002 (first entry)  
 DT  
 XX  
 DE Novel human diagnostic protein #3654.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PE 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Dmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB: AAS67850.  
 XX  
 XX  
 PS Claim 20; SEQ ID NO 34022; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 71 AA:  
 Query Match 2.8%; Score 8; DB 22; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 LGIGIALA 68  
 DB 16 LGIGIALA 23  
 RESULT 37  
 AAU61282  
 ID AAU61282 standard; Protein: 84 AA.  
 AC  
 XX AAU61282;  
 XX  
 DT 27-FEB-2002 (first entry)

XX  
 DE Propionibacterium acnes immunogenic protein #22178.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PE 20-APR-2001; 2001WO-US12865.  
 XX  
 PR 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI: 2001-616774/71.  
 DR N-PSDB: AAS59615.  
 XX  
 XX  
 PS Example 1; SEQ ID NO 22477; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 84 AA:  
 Query Match 2.8%; Score 8; DB 22; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 259 LRIRLPLPW 266  
 DB 34 LRIRLPLPW 41  
 RESULT 38  
 ABG20259  
 ID ABG20259 standard; Protein: 110 AA.  
 AC  
 XX ABG20259;  
 XX  
 DT 18-FEB-2002 (first entry)

```

XX DE Novel human diagnostic protein #20250.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HSE-) HXSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR N-PSDB: AAS84446.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PS biodiversity.
XX PS Claim 20; SEQ ID NO 50618; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations in
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 110 AA:
XX
XX Query Match 2.8%; Score 8; DB 22; Length 110;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 61 LGIGLALA 68
XX DB 44 LGIGLALA 51
XX
XX RESULT 39
XX AAO08094
XX ID AAO08094 standard; Protein; 117 AA.
XX AC AAO08094;
XX XX
XX XX 06-NOV-2001 (first entry)
XX DT
XX XX Human polypeptide SEQ ID NO 21986.
XX DE

```

```

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US04927.
XX PR 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA (HSE-) HXSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR N-PSDB: AA188025.
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders.
XX PS Claim 20; SEQ ID NO 21986; 1399pp + Sequence Listing; English.
XX PS
XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
XX CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 117 AA:
XX
XX Query Match 2.8%; Score 8; DB 22; Length 117;
XX Best Local Similarity 100.0%; Pred. No. 27;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 61 LGIGLALA 68
XX DB 54 LGIGLALA 61
XX
XX RESULT 40
XX ABB67712
XX ID ABB67712 standard; Protein; 184 AA.
XX AC ABB67712;
XX XX
XX XX 26-MAR-2002 (first entry)
XX DT
XX XX Drosophila melanogaster polypeptide SEQ ID NO 29928.
XX XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX XX
XX OS Drosophila melanogaster.
XX DT
XX XX
XX PN WO200171042-A2.
XX XX

```

PD 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PERE ) PE CORP NY.  
PA  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL11815.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PI interactions -  
XX  
XX Disclosure; SEQ ID NO 29928; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (AB57737-AB872072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 184 AA;  
  
Query Match 2.8%; Score 8; DB 22; Length 184;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 72 LLLAVVSL 79  
DB 16 LLLAVVSL 23  
  
RESULT 41  
ABP28041  
ID ABP28041 standard; Protein; 190 AA.  
XX  
XX ABP28041;  
AC  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Streptococcus polypeptide SEQ ID NO 5258.  
XX  
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX  
XX Streptococcus agalactiae.  
OS  
XX  
PN WO200234771-A2.  
PD  
XX  
PD 02-MAY-2002.  
XX  
XX 29-OCT-2001; 2001WO-GB04789.  
PF  
XX  
XX 27-OCT-2000; 2000GB-0026333.  
PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelin H;  
XX  
XX WPI; 2002-352536/38.  
DR N-PSDB; ABN68672.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PI for detecting a compound that binds to the protein -  
XX  
XX Claim 1; Page 3689; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
SQ Sequence 190 AA;  
  
Query Match 2.8%; Score 8; DB 23; Length 190;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 258 SLRIRTLR 265  
DB 50 SLRIRTLR 57  
  
RESULT 42  
AAM23684  
ID AAM23684 standard; Protein; 198 AA.  
XX  
XX AAM23684;  
AC  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Human EST encoded protein SEQ ID NO: 1209.  
XX  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX  
XX Homo sapiens.  
OS  
XX  
PN WO200154477-A2.  
PD  
XX  
PD 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02687.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;

DR WPI: 2001-476164/51.  
 DR N-PSDB: AAB98343.  
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use -  
 PS Claim 20; Page 873; 1275pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a  
 CC protein of the invention.  
 CC  
 SQ Sequence 198 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 54 ALVPLAL 61  
 |||||  
 DB 2 ALVPLAL 9  
 RESULT 43  
 ABP41674  
 ID ABP41674 standard; Protein: 222 AA.  
 AC ABP41674;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HPAMG11. SEQ ID NO:2806.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-US18569.  
 XX  
 PR 07-JUN-2000; 2000US-209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI: 2002-147878/19.  
 DR N-PSDB: ABQ54751.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases -  
 PS Claim 11; SEQ ID NO 2806; 2922pp; English.  
 XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 222 AA;

Query Match 2.8%; Score 8; DB 23; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLAL 67  
 |||||  
 DB 190 ALGGLAL 197

RESULT 44  
 ABG16279  
 ID ABG16279 standard; Protein: 286 AA.  
 AC ABG16279;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX

DE Novel human diagnostic protein #16270.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW  
 OS Homo sapiens.  
 XX

PN WO200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX

PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB: AAS80466.  
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 06:54:40 ; Search time: 46 Seconds

(without alignments)  
593.526 Million cell updates/sec

Title: US-09-245-198A-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR\_73:\*

1: PIR1:\*\n2: PIR2:\*\n3: PIR3:\*\n4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	2	A85866 hypothetical prote
2	10	3.2	733	2	S78376 photosystem I P700
3	8	2.8	58	2	A58208 protamine I-1 - pa
4	8	2.8	143	2	G84168 hypothetical prote
5	8	2.8	197	2	E72374 hypothetical prote
6	8	2.8	220	2	AG3547 bicyclomycin resis
7	8	2.8	278	2	D83080 hypothetical prote
8	8	2.8	339	2	C71132 hypothetical prote
9	8	2.8	372	2	H70813 probable cysteine
10	8	2.8	379	2	E64300 formate dehydrogen
11	8	2.8	381	2	AH3041 conserved hypotet
12	8	2.8	387	2	D84885 hypothetical prote
13	8	2.8	397	2	D98244 hypothetical prote
14	8	2.8	422	2	A60503 sperm-binding glyco
15	8	2.8	443	2	T17220 hypothetical prote
16	8	2.8	465	2	AC0347 probable membrane
17	8	2.8	471	2	A75267 gamma-ATP-ATP-bindi
18	8	2.8	576	2	E64186 gamma-glutamyltran
19	8	2.8	586	2	A41125 protamine II-3 - p
20	7	2.5	45	2	D58208
21	7	2.5	50	2	S22582
22	7	2.5	58	2	S34045
23	7	2.5	86	2	F87604
24	7	2.5	102	2	F87993
25	7	2.5	115	2	PH1560
26	7	2.5	115	2	H83201
27	7	2.5	118	1	IEEC5B
28	7	2.5	118	1	AE1753
29	7	2.5	123	2	AH2707 conserved hypotet

30	7	2.5	125	2	T27519	hypothetical prote
31	7	2.5	131	2	I52290	interleukin-13 - r
32	7	2.5	131	2	E30552	T-cell activation
33	7	2.5	146	2	T37116	probable transpos
34	7	2.5	147	2	A71217	hypothetical prote
35	7	2.5	150	2	T08734	hypothetical prote
36	7	2.5	157	2	S31078	seed allergen RA5
37	7	2.5	157	2	T02664	allergen - rice
38	7	2.5	157	2	A75567	conserved hypotet
39	7	2.5	157	2	E75350	hypothetical prote
40	7	2.5	160	2	S59925	allergen RA5B prec
41	7	2.5	161	1	DNEC17	outer membrane pro
42	7	2.5	161	2	D90651	histone-like prote
43	7	2.5	161	2	D85502	hypothetical prote
44	7	2.5	162	2	T24937	hypothetical prote
45	7	2.5	162	2	T31173	hypothetical prote

## ALIGNMENTS

RESULT 1  
A85866  
hypothetical protein Z3516 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85866  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma  
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoudis, K.; Apod  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: A85866  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Cross-references: 1-111 <STO>  
A:Cross-references: GB:AE005174; NID:q12516604; PIDN:AAG57389.1; GSPDB:GN00145; UWGP  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z3516

Query Match 3.5%; Score 10; DB 2; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.073;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGTAACTGCL 72  
DB 40 LGTAACTGCL 49  
|||||  
|||

RESULT 2  
S78376  
Photosystem I P700 apoprotein A2 - Odontella sinensis chloroplast  
C:Species: chloroplast Odontella sinensis  
C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 20-Jun-2000  
C:Accession: S78376  
R:Kowallik, K.V.; Stoebe, B.; Schafran, I.; Kroth-Panet, P.; Freier, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sin  
A:Reference number: S78238  
A:Accession: S78376  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-733 <KOW>  
A:Cross-references: EMBL:Z67753; NID:q1185127; PIDN:CAA91749.1; PID:q1185266  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 199  
C:Genetics:  
A:Gene: psab  
A:Genome: chloroplast  
C:Superfamily: photosystem I P700 apoprotein  
C:Keywords: chloroplast; electron transfer; membrane protein; membrane-associated co

Query Match 3.2%; Score 9; DB 2; Length 733;

Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 63 LGIALACIG 71  
|||||||  
Db 333 LGIALACIG 341

RESULT 3  
A58208  
protamine I-1 - painted turtle  
C:Species: Chrysemys picta (painted turtle)  
C>Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999  
C:Accession: A58208  
R:Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.  
J. Biol. Chem. 271, 23547-23557, 1996  
A:Title: Protamines of reptiles.  
A:Reference number: A58208; MID:96394458; PMID:8798564  
A:Accession: A58208  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-58 <HUN>  
C:Superfamily: sperm histone

Query Match 2.8%; Score 8; DB 2; Length 58;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 42 QRRGRG 49  
|||||||  
Db 35 QRRGRG 42

RESULT 4  
G84168  
hypothetical protein Vng0080h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84168  
R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.  
; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MID:20504483; PMID:11016950  
A:Accession: G84168  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-143 <STO>  
A:Cross-references: GB:AE004437; MID:g10579733; PIDN:AG18715.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG0080H

Query Match 2.8%; Score 8; DB 2; Length 143;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 61 LGIGLALA 68  
|||||||  
Db 55 LGIGLALA 62

RESULT 5  
E72374  
hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: E72374  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MID:99287316; PMID:10360571  
A:Accession: E72374  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <ARN>  
A:Cross-references: GB:AE001724; GB:AE000512; MID:g4980966; PIDN:AAD3553.1; PID:g498  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0469

Query Match 2.8%; Score 8; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 SLDDFEIS 9  
|||||||  
Db 135 SLDDFEIS 142

RESULT 6  
AG3547  
bicyclicomycin resistance protein [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AG3547  
R:Belveccio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Muter, C.; Ios, T.; Ivanov  
; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit  
A:Reference number: AD252; PMID:1175668  
A:Accession: AG3547  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AA153546.1; PID:g17984455; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10304  
A:Map position: 11

Query Match 2.8%; Score 8; DB 2; Length 220;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 54 ALLVPLAL 61  
|||||||  
Db 132 ALLVPLAL 139

RESULT 7  
D83080  
hypothetical protein PA4521 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83080  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: AB2950; MID:20437337; PMID:10984043  
A:Accession: D83080  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <STO>  
A:Cross-references: GB:AE004866; GB:AE004091; MID:g9950760; PIDN:AA07909.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4521

Query Match 2.8%; Score 8; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77  
 |||||  
 DB 47 LGLLAVV 54

RESULT 8  
 C71132  
 hypothetical protein PH0824 - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: C71132  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: C71132  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-339 <KAP>  
 A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29917.1; PID:g3257234  
 A:Experimental source: strain 073  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank C:Genetics:  
 A:Gene: PH0824  
 C:Superfamily: conserved hypothetical protein MTR900

Query Match 2.8%; Score 8; DB 2; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77  
 |||||  
 DB 111 LGLLAVV 118

RESULT 9  
 H70813  
 probable cysteine synthase - *Mycobacterium tuberculosis* (strain H37Rv)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: H70813  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: H70813  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-312 <COB>  
 A:Cross-references: GB:AL022004; GB:AL124345; NID:g3261550; PIDN:CAAL7654.1; PID:g291690  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: cysM3  
 C:Superfamily: threonine dehydratase

Query Match 2.8%; Score 8; DB 2; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 DB 96 LGIGLALA 103

RESULT 10  
 E64300  
 formate dehydrogenase (EC 1.2.1.2) beta chain - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: E64300  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Bla, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woes A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannas A:Reference number: A64300; MUID:96537999; PMID:8688087*  
 A:Accession: E64300  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-379 <BUR>  
 A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97986.1; PID:g159081 C:Genetics:  
 A:Map position: REV7250-6111  
 C:Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homology  
 C:Keywords: oxidoreductase  
 F:273-348/Domain: ferredoxin 2[4Fe-4S] homology <FE>

Query Match 2.8%; Score 8; DB 2; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDGVVAL 224  
 |||||  
 DB 35 LVDGVVAL 42

RESULT 11  
 AH3041  
 conserved hypothetical protein Atu3948 [imported] - *Agrobacterium tumefaciens* (strain AH3041)  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AH3041  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Mc erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; MCC ; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58. A:Reference number: AB2577; PMID:11743193  
 A:Accession: AH3041  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-381 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AAL44750.1; PID:g17742385; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu3948  
 A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258  
 |||||  
 DB 325 LALRPGSS 332

RESULT 12  
 D84885  
 hypothetical protein At2g45000 [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cross)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84885  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84885  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-387 <STO>  
A:Cross-references: GB:AE002093; NID:94895250; PIDN:AAD32835.1; GSPDB:GN00139  
A:Gene: At2g45000  
A:Map position: 2

Query Match 2.8%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 EELVAEED 100  
DB 370 EELVAEED 377

## RESULT 13

D98244  
hypothetical protein AGR\_L1808 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: D98244  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; PMID:11743194  
A:Accession: D98244  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:q15159347; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L1808  
A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRGS 258  
DB 341 LALRGS 348

## RESULT 14

A60503  
sperm-binding glycoprotein ZP3 precursor - golden hamster  
N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3  
C:Species: *Mesocricetus auratus* (golden hamster)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A60503  
R:Kunlich, R.A.; Ruiz-Seller, B.; Massarman, P.M.  
Dev. Biol. 142, 414-421, 1990  
A:Title: Genomic organization and polypeptide primary structure of zona pellucida glycop  
A:Reference number: A60503; MUID:91078540; PMID:2257975  
A:Accession: A60503  
A:Molecule type: DNA  
A:Residues: 1-422 <KIN>

A:Cross-references: GB:M63629  
A:Note: the authors translated the codon CAA for residue 251 as Glu, and AGG for residue  
C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor  
C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology  
C:Keywords: glycoprotein; oocyte  
F:45-300/Domain: ZP domain homology <ZPH>

Query Match 2.8%; Score 8; DB 1; Length 422;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLA 66  
DB 386 LALGLA 393

## RESULT 15

T17220  
hypothetical protein DKFZp566001.1 - human

C:Species: *Homo sapiens* (man)  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Jun-2000  
C:Accession: T17220  
R:Bium, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A:Reference number: Z18725  
A:Accession: T17220  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-443 <BLU>

A:Cross-references: EMBL:AL117414  
A:Experimental source: fetal kidney; clone DKFZp566001  
C:Genetics:  
A:Note: DKFZp566001.1  
C:Superfamily: gamma-glutamyltransferase

Query Match 2.8%; Score 8; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLA 68  
DB 14 LGGLA 21

## RESULT 16

AC0347  
probable membrane protein yegB [imported] - *Yersinia pestis* (strain CO92)

C:Species: *Yersinia pestis*  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 17-May-2002  
C:Accession: AC0347  
R:Parikh, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.  
deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0347  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-465 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92102.1; PID:q15980820; GSPDB:GN00175  
C:Genetics:  
A:Gene: yegB  
C:Superfamily: multidrug-efflux transporter

Query Match 2.8%; Score 8; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLAWSL 79  
DB 335 LLAWSL 342

## RESULT 17

A75267  
probable transport protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: A75267  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
M.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999.  
 A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: A75267  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-471 <WH>  
 A:Cross-references: GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAFI2043.1; PID:g646032  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2502  
 A:Map position: 1

Query Match 2.8%; Score 8; DB 2; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGIAL 67  
 DB 366 ALGIAL 373

RESULT 18  
 E64186  
 A:Title: ATP-binding transport protein H1156 - *Haemophilus influenzae* (strain Rd KW20)  
 C:Species: *Haemophilus influenzae*  
 C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 02-Feb-2001  
 C:Accession: E64186  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Goehayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.  
 A:Reference number: A64000; MUID:9550630; PMID:7542800  
 A:Accession: E64186  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-576 <TIGR>  
 A:Cross-references: GB:U3795; GB:LA2023; NID:g1574708; PIDN:AAC2281.1; PID:g1574713;  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP; nucleotide binding; P-loop  
 F:355-550/Domain: ATP-binding cassette homology <ABC>  
 F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match 2.8%; Score 8; DB 2; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 PLALGIAL 65  
 DB 159 PLALGIAL 166

RESULT 19  
 A41125  
 A:Title: gamma-glutamyltransferase (EC 2.3.2.2) related protein - human  
 N:Alternate names: gamma-glutamyltransferase-like activity 1; GGT-REL  
 C:Species: *Homo sapiens* (man)  
 C>Date: 27-Mar-1992 #sequence\_revision 27-Mar-1992 #text\_change 18-Jun-1999  
 C:Accession: A41125  
 R:Heisterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Groffen, J. Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991  
 A:Title: Identification of a human gamma-glutamyl cleaving enzyme related to, but distinct from, rat gamma-glutamyl transferase  
 A:Reference number: A41125; MUID:91296809; PMID:1676842  
 A:Accession: A41125  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-586 <HEI>  
 A:Cross-references: GB:M64099; NID:g183141; PIDN:AA58503.1; PID:g183142  
 C:Genetics:  
 A:Gene: GDB:GGTAL1; GGT-REL  
 A:Cross-references: GDB:134033

C:Superfamily: gamma-glutamyltransferase  
 C:Keywords: aminoacyltransferase; glycoprotein; transmembrane protein

Query Match 2.8%; Score 8; DB 2; Length 586;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68  
 DB 14 LGIGIALA 21

RESULT 20  
 D58208  
 A:Title: Protamines of reptiles.  
 C:Species: *Chrysemys picta* (painted turtle)  
 C>Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999  
 C:Accession: D58208  
 R:Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp J. Biol. Chem. 271, 23547-23557, 1996  
 A:Title: Protamines of reptiles.  
 A:Reference number: A58208; MUID:96394458; PMID:8798564  
 A:Accession: D58208  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-45 <HUN>  
 C:Superfamily: sperm histone

Query Match 2.5%; Score 7; DB 2; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 DB 24 RRRGRG 30

RESULT 21  
 S22582  
 A:Title: Protamine 1 - *Saguinus imperator*  
 C:Species: *Saguinus imperator*  
 C>Date: 22-Nov-1993 #sequence\_revision 12-Apr-1996 #text\_change 21-Jul-2000  
 C:Accession: S22582  
 R:Queralto, R.; Oliva, R. Nucleic Acids Res. 19, 5786, 1991  
 A:Title: Protamine 1 gene sequence from the primate *Saguinus imperator* isolated with  
 A:Reference number: S22582; MUID:92051332; PMID:1840669  
 A:Accession: S22582  
 A:Molecule type: DNA  
 A:Residues: 1-50 <QUE>  
 A:Cross-references: EMBL:X61678; NID:g58405; PIDN:CAA43853.1; PID:g4494091  
 A>Note: the authors translated the codon TAC for residue 43 as Thr  
 C:Genetics:  
 A:Introns: 37/1  
 C:Superfamily: sperm histone  
 C:Keywords: chromosomal protein; DNA binding; nucleus; spermatogenesis

Query Match 2.5%; Score 7; DB 2; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRGRR 48  
 DB 18 ORRGRR 24

RESULT 22  
 S34045  
 A:Title: North American opossum  
 C:Species: *Didelphis virginiana*, *Didelphis marsupialis virginiana* (North American OP  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
 C:Accession: S34045  
 R:Winkfein, R.J.; Nishikawa, S.; Connor, W.; Dixon, G.H.

Eur. J. Biochem. 215, 63-72, 1993  
 A:Title: Characterization of a marnsupial sperm protamine gene and its transcripts from  
 A:Reference number: S34045; MUID:93345500; PMID:8344286  
 A:Accession: S34045  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-58 <MIN>  
 A:Cross-references: EMBL:X74044; NID:9407062; PIDN:CAAS2193.1; PID:9407063  
 C:Superfamily: sperm histone  
 C:Keywords: DNA binding; nucleus

Query Match 2.5%; Score 7; DB 2; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 |||||  
 Db 35 RRRGRG 41

## RESULT 23

F87604  
 hypothetical protein CC2870 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: F87604  
 R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Lamb, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: F87604  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-86 <STO>  
 A:Cross-references: GB:AE005673; NID:913424486; PIDN:AAK24834.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2870

Query Match 2.5%; Score 7; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGL 65  
 |||||  
 Db 26 LALGLGL 32

## RESULT 24

F87993  
 protein ZC334.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

C:Accession: F87993  
 R:anonymous, The C. elegans Sequencing Consortium.

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
 Science 282, 2012-2018, 1998  
 A:Reference number: A75000; MUID:99069618; PMID:9851916

A:Note: see webistes genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: F87993

A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-102 <STO>  
 A:Cross-references: GB:chr\_I; PIDN:CAH04964.1; PID:93881432; GSPDB:GN00019; CESP:ZC334.3  
 A:Note: predicted using GeneFinder

C:Genetics:  
 A:Gene: ZC334.3  
 A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 55 LVLPLAL 61  
 |||||  
 Db 33 LVLPLAL 39

## RESULT 25

PH1560

Ig heavy chain V region (clone VH32) - human (fragment)

C:Species: Homo sapiens (man)  
 C:Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C:Accession: PH1560  
 R:Kassam, L.Z.; Kipp, T.J.

J. Exp. Med. 177, 1039-1046, 1993

A:Reference number: PH1557; MUID:93210459; PMID:7681468  
 A:Accession: PH1560

A:Molecule type: DNA

A:Residues: 1-115 <RAS>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

F:32-115/Domain: Immunoglobulin homology <IMM>

Query Match 2.5%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 76  
 |||||  
 Db 5 LGLLAV 11

## RESULT 26

H83201

conserved hypothetical protein PA357 [imported] - Pseudomonas aeruginosa (strain PAC

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: H83201  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; I

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: H83201

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-115 <STO>

A:Cross-references: GB:AE004776; GB:AE004091; NID:99949701; PIDN:AA06945.1; GSPDB:GT

A:Experimental source: strain PAO1

C:Genetics:  
 A:Gene: PA357

Query Match 2.5%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGL 65  
 |||||  
 Db 49 LALGLGL 55

RESULT 27  
 IEE5B  
 hypothetical protein, 12K - Escherichia coli insertion sequence ISS

C:Species: Escherichia coli

C:Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 10-Sep-1999

C:Accession: B91483; C03582; A04466

R:Schoner, B.; Kahn, M.

Gene 14, 165-174, 1981

A:Title: The nucleotide sequence of ISS from Escherichia coli.

A:Reference number: A91483; MUID:82028653; PMID:6269959



A:Accession: B91483  
 A:Molecule type: DNA  
 A:Residues: 1-118 <SCH>  
 R:Engler, J.A.; van Bree, M.P.  
 Gene 14, 155-163, 1981  
 A:Title: The nucleotide sequence and protein-coding capability of the transposable element  
 A:Reference number: A03582; MUID:82028652; PMID:6269958  
 A:Accession: C03582  
 A:Molecule type: DNA  
 A:Residues: 1-118 <ENG>  
 C:Genetics:  
 A:Mobile element: insertion sequence IS5  
 C:Superfamily: Escherichia coli hypothetical protein o263

Query Match 2.5%; Score 7; DB 1; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DGGAVRQ 29  
 |||||  
 Db 107 DGGAVRQ 113

RESULT 28  
 AE1753  
 C:Species: bacteriophage b1285 homolog 11n2570 [imported] - Listeria innocua (strain Clp1  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 A:Accession: AE1753  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fshl, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Me  
 Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AE1753  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-118 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC97797.1; PID:q16415092; GSPDB:GN00178  
 A:Experimental source: strain Clp11262  
 C:Genetics:  
 A:Gene: 11n2570

Query Match 2.5%; Score 7; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLDFEI 8  
 |||||  
 Db 61 SLDFEI 67

RESULT 29  
 AH2707  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 A:Accession: AH2707  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Moks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McCell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Teo, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AH2707  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-123 <KOR>

A:Cross-references: GB:AE008688; PIDN:AA142078.1; PID:q17739458; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atul065  
 A:Map position: circular chromosome

Query Match 2.5%; Score 7; DB 2; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RLPLRS 18  
 |||||  
 Db 50 RLPLRS 56

RESULT 30  
 T27519  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 A:Accession: T27519  
 R:McLay, K.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z20381  
 A:Accession: T27519  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-125 <WIL>  
 A:Cross-references: EMBL:820882; PIDN:CAB04964.2; GSPDB:GN00019; CESP:ZC334.3  
 A:Experimental source: clone ZC334  
 C:Genetics:  
 A:Gene: CESP:ZC334.3  
 A:Map position: 1  
 A:introns: 99/1

Query Match 2.5%; Score 7; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LVLPLAL 61  
 |||||  
 Db 33 LVLPLAL 39

RESULT 31  
 I52290  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
 A:Accession: I52290  
 R:Lakkis, F.G.; Gruet, E.N.  
 Biochem. Biophys. Res. Commun. 197, 612-618, 1993  
 A:Title: Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expre  
 A:Reference number: I52290; MUID:94092138; PMID:791615  
 A:Accession: I52290  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-131 <RES>  
 A:Cross-references: GB:L26913; NID:9438875; PIDN:AAA16478.1; PID:9438876  
 C:Genetics:  
 A:Gene: IL-13  
 C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACG 71  
 |||||  
 Db 9 LALACG 15

RESULT 32

E30552  
T-cell activation protein P600 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 16-Jul-1999  
C:Accession: E30552  
R:Brown, K.D.; Zurawski, S.M.; Mosmann, T.R.; Zurawski, G.  
J. Immunol. 142, 679-687, 1989  
A:Title: A family of small inducible proteins secreted by leukocytes are members of a new class of various activation processes.  
A:Reference number: A30552; MUID:89093958; PMID:2521353  
A:Accession: E30552  
A:Molecule type: mRNA  
A:Residues: 1-131 <BRO>  
A:Cross-references: GB:M23504; NID:9533246; PIDN:AAA40149.1; PID:9533247  
C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACIG 71  
|||||||  
DB 9 LALACIG 15

RESULT 33  
T37116  
probable transposase, truncated [imported] - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000  
C:Accession: T37116  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M. submitted to the EMBL Data Library, August 1999  
A:Reference number: Z31588  
A:Accession: T37116  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-146 <SAU>  
A:Cross-references: EMBL:AL109950; PIDN:CAB52967.1; GSPDB:GN00070; SCOEDB:SCJ4.33c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCJ4.33c  
C:Superfamily: Synchocystis transposase s111710

Query Match 2.5%; Score 7; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLIAY 76  
|||||||  
DB 102 LGLLIAY 108

RESULT 34  
A71217  
hypothetical protein PH2001 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: A71217  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekir M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: A71217  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-147 <KAW>  
A:Cross-references: GB:AP000007; GB:AP000001; NID:93236134; NID:93236128; PID:93258445; A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank A:Note: this sequence is split into two separate translations in GenBank (residues 1-124 C:Genetics:

A:Gene: PH2001  
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001

Query Match 2.5%; Score 7; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LACIGLL 73  
|||||||  
DB 45 LACIGLL 51

RESULT 35  
T08734  
hypothetical protein DKFZp566F0546.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999  
C:Accession: T08734  
R:Ottenwelder, B.; Obermayer, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S. submitted to the Protein Sequence Database, May 1999  
A:Reference number: Z16474  
A:Accession: T08734  
A:Molecule type: mRNA  
A:Residues: 1-150 <OTT>  
A:Cross-references: EMBL:AL050075  
A:Experimental source: fetal kidney; clone DKFZp566F0546  
C:Genetics:  
A:Note: DKFZp566F0546.1

Query Match 2.5%; Score 7; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AASSLGP 240  
|||||||  
DB 32 AASSLGP 38

RESULT 36  
S31078  
seed allergen RA5 - rice  
C:Species: Oryza sativa (rice)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: S31078  
R:Adachi, T.; Izumi, H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda, Plant Mol. Biol. 21, 239-248, 1993  
A:Title: Gene structure and expression of rice seed allergenic proteins belonging to A:Reference number: S31078; MUID:93144699; PMID:7678765  
A:Accession: S31078  
A:Molecule type: mRNA  
A:Residues: 1-157 <ADA>  
A:Cross-references: EMBL:D11430; NID:9218196; PIDN:BAA01966.1; PID:9218197  
C:Superfamily: wheat alpha-amylase inhibitor  
C:Keywords: seed

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVS 78  
|||||||  
DB 11 LLLAVVS 17

RESULT 37  
T02664  
allergen - rice  
C:Species: Oryza sativa (rice)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Jul-1999  
C:Accession: T02664  
R:Yun, C.H.; Park, J.H.; Eun, M.Y. submitted to the EMBL Data Library, January 1998  
A:Description: Nucleotide sequence of rice allergenic protein.

A:Reference number: 214691  
 A:Accession: F02664  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-157 <YUN>  
 A:Cross-references: EMBL:AF042200; NID:g2827315; PIDN:AAB99797.1; PID:g2827316  
 A:Experimental source: strain N1pponbare  
 C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78  
 |||||  
 DB 11 LLLAVS 17

## RESULT 38

A75567

conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

A:Accession: A75567

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75567

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 &lt;WHI&gt;

A:Cross-references: GB:AE001867; GB:AE000513; NID:96457693; PIDN:AAF09626.1; PID:9645769

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0033

A:Map position: 1

Query Match

Best Local Similarity 2.5%; Score 7; DB 2; Length 157;  
 100.0%; Pred. No. 75;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 IRTLPWA 267  
 |||||  
 DB 22 IRTLPWA 28

## RESULT 39

E75530

hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

A:Accession: E75530

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75530

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 &lt;WHI&gt;

A:Cross-references: GB:AE001865; GB:AE000513; NID:96458024; PIDN:AAF09935.1; PID:9645803

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0352

A:Map position: 1

Query Match

Best Local Similarity 2.5%; Score 7; DB 2; Length 157;  
 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 |||||  
 DB 144 RRRGRG 150

## RESULT 40

S59925

allergen RA5B precursor - rice

C:Species: *Oryza sativa* (rice)

C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Jun-2000

A:Accession: S59925

R:Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsuda, T.

Biochim. Biophys. Acta 1251, 201-204, 1995

A:Title: Classification of rice allergenic protein cDNAs belonging to the alpha-amyl

A:Reference number: S59925

A:Accession: S59925

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-160 &lt;AIY&gt;

A:Cross-references: EMBL:D42142; NID:g1398917; PIDN:BA07713.1; PID:g1398918

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78  
 |||||  
 DB 11 LLLAVS 17

## RESULT 41

DNEC17

outer membrane protein h1pa precursor - *Escherichia coli* (strain K-12)

N:Alternate names: DNA-binding 17K protein; histone-like protein hlp

C:Species: *Escherichia coli*  
 C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 01-Mar-2002

A:Accession: J70304; A38063; S13728; B64742; I54944; S20426

R:Holck, A.; Klepe, K.

Gene 67, 117-124, 1988

A:Title: Cloning and sequencing of the gene for the DNA-binding 17K protein of *Esche*

A:Reference number: J70304; MUID:88329735; PMID:2843433

A:Accession: J70304

A:Molecule type: DNA

A:Residues: 1-161 &lt;HOI&gt;

A:Cross-references: GB:M21118; NID:9147821; PIDN:AAA24630.1; PID:9147822

A:Accession: A38063

A:Molecule type: protein

A:Residues: 21-30 &lt;HO2&gt;

A:Experimental source: strain B

R:Dickler, I.B.; Seetharam, S.

J. Bacteriol. 173, 334-344, 1991

A:Title: Cloning and nucleotide sequence of the *fira* gene and the *fira200*(Ts) allele

A:Reference number: S13728; MUID:91100302; PMID:1987124

A:Accession: S13728

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 72-161 &lt;DIC&gt;

A:Cross-references: EMBL:X54797; NID:941468; PIDN:CAA38567.1; PID:941469

A:Experimental source: strain K-12, substrain MG1655

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Petina, N.T.; Burland, V.; Riley, M.

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: B64742

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-161 &lt;BLAT&gt;

A:Cross-references: GB:AE000127; GB:U00096; NID:g1786370; PIDN:AAC73289.1; PID:g1786

A:Experimental source: strain K-12, substrain MG1655

R.Hirvas, L.; Koski, P.; Vaara, M.  
 J. Bacteriol. 173, 1223-1229, 1991  
 A:Title: The ompH gene of *Yersinia enterocolitica*: cloning, sequencing, expression, and  
 A:Reference number: 154944; MUID:91123198; PMID:1991717  
 A:Accession: 154944  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-14, 'L', 16-148, 'E', 150-152, 'I', 154-161 <RES>  
 A:Cross-references: EMBL:X75465; NID:9432661; PIDN:CA53207.1; PID:9432662  
 C:Comment: The h1pa protein has been believed to be a histone-like constituent of bacteria  
 C:Genetics:  
 A:Gene: h1pa, sfp  
 C:Superfamily: DNA-binding 17K protein  
 C:Keywords: membrane protein  
 F:1-20/Domain: signal sequence  
 F:21-161/Product: outer membrane protein h1pa #status predicted <STG>

Query Match 2.5%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 42  
 D90651  
 histone-like protein H1pa [imported] - *Escherichia coli* (strain O157:H7, substrain RMD  
 C:Species: *Escherichia coli*  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 A:Accession: D90651  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 gasevaya, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shihagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene  
 A:Reference number: A95629; MUID:21156231; PMID:11258796  
 A:Accession: D90651  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA833603.1; PID:913359636; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: EC80180  
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 43  
 D85502  
 hypothetical protein h1pa [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 A:Accession: D85502  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Diallanita, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: D85502  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <SNO>  
 A:Cross-references: GB:AE005174; NID:912512906; PIDN:AAG54480.1; GSPDB:GN00145; UWGP:201  
 A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:  
 A:Gene: h1pa  
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 44  
 T24937  
 hypothetical protein W03C9.4 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 A:Accession: T24937; T26123  
 R:Gregory, J.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: T24937  
 A:Accession: T24937  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <WIL>  
 A:Cross-references: EMBL:Z50015; PIDN:CAA90314.1; GSPDB:GN00020; CESP:W03C9.4  
 A:Experimental source: clone T15G9  
 R:Gregory, J.; Alnscough, R.  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z20155  
 A:Accession: T26123  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <WIZ>  
 A:Cross-references: EMBL:Z66516; PIDN:CAA91361.1; GSPDB:GN00020; CESP:W03C9.4  
 A:Experimental source: clone W03C9  
 C:Genetics:  
 A:Gene: CESP:W03C9.4  
 A:Map position: 2  
 A:introns: 16/3; 64/1; 120/3; 128/1

Query Match 2.5%; Score 7; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 231 SATASS 237  
 |||||  
 DB 112 SATASS 118

RESULT 45  
 T31173  
 hypothetical protein 424 - *Sphingomonas aromaticivorans* plasmid pNLI  
 C:Species: *Sphingomonas aromaticivorans*  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jun-2000  
 A:Accession: T31173  
 R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Slak, E.C.; Sensen, C.W.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: Complete sequence of a 184 kb catabolic plasmid from *Sphingomonas* *aro*  
 A:Reference number: Z20992  
 A:Accession: T31173  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <ROM>  
 A:Cross-references: EMBL:AF079317; NID:93378261; PID:93378314; PIDN:AAD03897.1  
 C:Genetics:  
 A:Genome: plasmid pNLI  
 A:Note: ori424  
 C:Superfamily: *Sphingomonas aromaticivorans* hypothetical protein 424  
 Query Match 2.5%; Score 7; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 77;

	Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	138	ARRATA	144							
Db	28	ARRATA	34							

Search completed: May 8, 2003, 07:07:55  
Job time : 50 secs

---

**THIS PAGE BLANK (USPTO)**

**From:** Chan, Christina  
**Sent:** Monday, May 05, 2003 3:41 PM  
**To:** Schnizer, Richard; STIC-Biotech/ChemLib  
**Subject:** RE: 09/245,198

**Please rush. Thanks Chris**

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

**From:** Schnizer, Richard  
**Sent:** Monday, May 05, 2003 3:21 PM  
**To:** Chan, Christina  
**Subject:** 09/245,198

Please authorize the following RUSH search for 09/245,198 which is an amended case in condition for allowance.

Please search the commercial and interference databases for:

-SEQ ID NOS:1-4, and

-nucleic acids that could encode the amino acid sequences of SEQ ID NOS:2 and 4

Please perform oligo searches on SEQ ID NOS: 3 and 4.

Thank you-

Richard Schnizer, Ph.D.  
Patent Examiner  
Art Unit 1635  
CM1 12E17  
703-306-5441  
Mail Box CM1 11E12

RECEIVED  
MAY - 5 2003  
STIC/CHEM DIVISION

Point of Contact:  
Thomas G. Larson, Ph.D.  
703-308-7309  
CM1, Rm. 6 B 01

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 5/6  
Date Completed: 5/12  
Searcher Prep/Review: 10  
Clerical: 10  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 3  
AA Sequences: 2  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**